6-1-2009

Monitoring with Collective Memory: Forgiveness for Optimally Empty Promises

David A. Miller

Kareen Rozen

MONITORING WITH COLLECTIVE MEMORY:
Forgiveness for Optimally Empty Promises

By

David A. Miller and Kareen Rozen

June 2009
Revised April 2010

COWLES FOUNDATION DISCUSSION PAPER NO. 1698

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Monitoring with collective memory: Forgiveness for optimally empty promises*

David A. Miller† UCSD

Kareen Rozen‡ Yale

April 20, 2010

Abstract

We study optimal contracting in a team setting with moral hazard, where teammates promise to complete socially efficient but costly tasks. Teammates must monitor each other to provide incentives, but each team member has limited capacity to allocate between monitoring and productive tasks. Players incur contractual punishments for unfulfilled promises that are discovered. We show that optimal contracts are generally "forgiving" and players optimally make "empty promises" that they don't necessarily intend to fulfill. As uncertainty in task completion increases, players optimally make more empty promises but fewer total promises. A principal who hires a team of agents optimally implements a similar contract, with profit-sharing and employment-at-will. When agents differ in their productivity, the model suggests a "Dilbert principle" of supervision: less productive players optimally specialize in monitoring the more productive players' promises.

Keywords: Teams, contracting, capacity constraints, empty promises, forgiveness, bounded memory, costly monitoring, moral hazard, Dilbert principle.

JEL Codes: C72, D03, D86.

---

*This paper originally circulated under the title "Need I remind you? Monitoring with collective memory." We are grateful to Larry Samuelson and Joel Sobel, as well as to Itzhak Gilboa, David Levine, Ben Polak, Andy Postlewaite, Ron Siegel, Andy Skrzypacz, Ana de la O Torres, Juuso Välimäki, and Bob Wilson for helpful comments and suggestions. This paper benefitted from the feedback of seminar participants at Columbia, London Business School, LSE, Penn State, Princeton, Queen Mary, Stanford GSB, Toronto, Yale, Zurich, the NSF/NBER/CEME Conference on General Equilibrium and Mathematical Economics, and the NAWM of the Econometric Society. Max Dvorkin and Jung-Myun Moon provided excellent research assistance.

†Address: Dept. of Economics, 9500 Gilman Dr. #0508, La Jolla, CA 92093–0508. E-mail: d9miller@ucsd.edu. Home page: http://dss.ucsd.edu/~d9miller. I thank Yale University and the Cowles Foundation for Research in Economics for hospitality and financial support.

‡Address: Dept. of Economics and the Cowles Foundation for Research in Economics, 30 Hillhouse Ave., New Haven, CT 06511. E-mail: kareen.rozen@yale.edu. Home page: http://www.econ.yale.edu/~kr287/. I thank the NSF for generous financial support through grant SES-0919955.

1 Introduction

In a team setting with moral hazard, teammates may need to monitor each other to motivate performance. However, each team member may have limited capacity to allocate between monitoring and productive tasks. Such resource constraints may arise, for example, from limited time, staffing, capital, attention, or memory. We propose a model in which teammates promise to complete socially efficient tasks—each of which is a single-person activity that requires costly effort to complete, but can be "botched" effortlessly. Monitoring does not require effort, but carries the opportunity cost of not being able to perform an additional task. The team commits to a schedule of punishments that punishes each player based on the number of her unfulfilled promises that are discovered; we call this a *counting contract*. We study optimal counting contracts and equilibrium behavior in this setting. Three main features drive our results:

- First, the capacity constraint leads to a tradeoff between production and monitoring, and leads to incomplete monitoring at the optimum.
- Second, each player privately learns which of his tasks are feasible, and then privately decides into which of his feasible tasks he will put effort. The monitoring technology allows his teammates to discover, with some probability, whether a task was uncompleted, but not whether an uncompleted task was feasible. This implies that punishments must be incurred with positive probability.
- Third, we assume punishments destroy surplus. That is, punishment takes the form of embarrassment, loss of status, or some other penalty that does not enrich one's teammates. This attempts to capture settings in which output-contingent transfers are not contractible, since using relational mechanisms to endogenously enforce transfers would crowd out the enforcement of other desired noncontractible behavior.

Since punishments are costly and occur with positive probability, the optimal incentive scheme should use them only sparingly. However, at the same time, resources allocated to increase monitoring—enabling a finer, attenuated punishment scheme—come at the expense of productive activity.[1]

Under an optimal contract, how should capacity be allocated between productive tasks and monitoring? How many tasks should teammates promise to complete? How many of those promises

[1] If the feasibility of a task could be publicly observed, the moral hazard problem would be trivial and punishments would not occur on the equilibrium path. The problem would also be uninteresting if the players could simply transfer utility. For example, with three or more players it would be possible to implement costless punishments, by rewarding a third player when one player's unfulfilled promise is discovered by a second player. With costless punishments, the players would put all but a minimal portion of their memory resources toward promise making. Indeed, if they could randomize, then there would be no well-defined optimum.

should they fulfill? And how do these change with the degree of uncertainty in production and monitoring? These questions lie at the intersection of the literatures on teams (e.g., Holmström 1982), contracting with costly monitoring (e.g., Williamson 1987, Border and Sobel 1987, Mookherhjee and Png 1989), public goods (e.g., Palfrey and Rosenthal 1984) and bounded rationality (e.g., Rubinstein 1998).[2]

Under an optimal contract, we find that players make "empty promises" they don't necessarily intend to fulfill, and are "forgiven" for having done so. When tasks are less likely to be feasible, the team members optimally make more empty promises, even as they make fewer promises in total. This suggests, for example, that workers in research and development—in which task feasibility is highly uncertain, and significant time commitments are required—should be assigned more responsibilities than they would actually fulfill, and that they should not be punished unless their observed failures exceed some threshold.

Our model applies to team settings where the collective ability to perform and monitor tasks is resource constrained. This is likely to be the case when each worker is a specialist and tasks are complex, as it may take another specialist to monitor him. According to Lazear and Shaw (2007), from 1987 to 1996 "the percent of large firms with workers in self-managed work teams rose from 27 percent to 78 percent," and moreover, "the firms that use teams the most are those that have complex problems to solve." Interpretations of the constraints facing a team may be grouped into two main categories. In a "classical" interpretation, the limiting resource is tangible, such as time, staffing, or capital. For example, to determine whether his teammate has satisfactorily completed a report, a member of a consulting team must spend time reading that report, rather than researching and writing his own report. Viewed in this way, we study optimal contracts in a setting with moral hazard, capacity constraints, private information (as to the number of tasks that may be completed), imperfect monitoring, and costly punishments. In a "bounded rationality" interpretation, the resource may be an aspect of conscious capacity that bounds the agents' rationality, such as cognitive ability, attention, or memory. Evidence from psychology suggests working memory is sharply bounded and imperfect. If a task is too complex to be fully specified in a convenient written form (such as a contract), an agent who promises to perform the task must expend memory resources to store the relevant details. Moreover, to be able to detect whether he has completed the task properly, another teammate must also hold these details in memory. If an agent forgets the details he stored, or decides to shirk by ignoring them, he must hope that his teammates will forget as well. We henceforth discuss the model and results in the language of bounded memory, identifying capacity with memory size and uncertainty in production and monitoring with imperfect recall (or forgetfulness).

[2]We view a contract as an informal agreement that is enforced by selecting among equilibria in some unspecified continuation game. In such a context, any common knowledge event at the end of the game is "contractible." A variety of related questions arise in the principal-agent literature. At the most basic level, we build on the seminal results on optimal contracts, such as Mirrlees (1999) and Holmström (1979).

1.1 Empty promises, heterogeneity, and organizational implications

Our model and the problem of designing an optimal contract are formalized in Sections 2 and 3, respectively. In Section 4, we study the benchmark case of linear contracts with bounded punishments. Because a linear contract treats each task separately and symmetrically, each player optimally completes as many promises as he can. However, we show in Section 5 that linear contracts are efficient only when the probability of recall is either very high or very low. For intermediate probabilities of recall, the optimal contract is generally nonlinear and *forgiving*: a player who fails to fulfill a small number of promises is not punished, making her unwilling to fulfill all her promises even if she can. Therefore, players make *empty promises*—promises they do not necessarily intend to fulfill. The empty-promises effect can be large: there is a range of parameters where making and keeping the maximal number of promises gives positive social utility, but is socially dominated by making half as many promises and keeping only one of these. Intuitively, empty promises buffer against the likelihood that some tasks will not be feasible. The corresponding contract must then leave enough monitoring slots to be able to forgive some failures; therefore players will "under-deliver" on their promises whenever it is optimal to "under-promise" relative to their memory.

We show that the optimal contract forgives up to a threshold and punishes linearly thereafter, either when either (i) the contract space is unrestricted but memory is tightly bounded (at most five promises), or (ii) the memory bound is arbitrary but contracts must deliver increasingly large punishments for larger numbers of unfulfilled promises. In each of these cases, the optimal contract induces each player to use a *cutoff strategy*, where she completes only up to a certain number of tasks out of the promises she recalls. For the special case of maximally forgiving contracts, which forgive all but the worst outcome (and which are fully optimal in case (i)), as players become more forgetful they optimally lower their cutoff for task completion and make fewer promises, raising their number of empty promises but allowing themselves to devote more of their memories to monitoring.

We allow players to differ on two dimensions of quality in Section 6: the number of tasks they can handle (memory capacity), and the amount of uncertainty they introduce in performing and monitoring tasks (forgetfulness). To focus on the endogenous allocation of supervisory responsibility, we examine linear contracts, under which empty promises do not arise (this restriction is without loss of generality when the probability of recall is high). When team members are asymmetric, we show that greater supervisory responsibility is optimally assigned to the less able player. Moreover, an increase in the strength of one player's memory reduces the number of tasks that her teammate optimally promises. This accords with the "Dilbert principle," which suggests that less productive team members should be removed from productive tasks (Adams 1996). Lazear (2004) argues that in an incomplete information setting, the related "Peter Principle" of Peter and Hull (1969)—which says that individuals are promoted to their point of incompetence—is a statistical necessity. By contrast, in a complete information setting we argue that when capacity is bounded, the less able worker is most effectively used to monitor the more able worker's task completion.

In Section 7 we consider the implications of our results for organizational structure, in the context of two stylized extensions to the model. In the first extension, each "player" in the original model is conceptualized as a team of several players, for which the size of the team constitutes the capacity constraint. Since the cost of effort accrues to a single player rather than to the group, non-linear contracts offer no benefit over linear contracts. In the second extension, we show that a principal who hires a team of agents and asks them to monitor each other optimally offers a forgiving contract using dilute profit-sharing, augmented by a rule of firing individual agents only for many observed failures.

Contracts of such a forgiving nature appear prevalent in the workforce, where incentive schemes often consist of a fixed wage, little or no bonus, and employment-at-will. In their survey, Baker, Jensen and Murphy (1987) find that "transitory performance-based bonuses seldom account for an important part of a worker's compensation." These facts suggest that despite their theoretical benefits, finely varying monetary incentive schemes and individualized bonuses are rarely observed in practice. Oyer and Schaefer (2005) find that many firms give broad-based stock options as compensation, but that these options have little incentive effect. Instead, their primary purpose seems to be employee retention. In our principal's optimal contract, profit-sharing plays a similar role—the principal has to satisfy the agents' participation constraints, and uses profit-sharing to just meet that constraint.

Mirrlees (1999) studies a principal-agent problem with moral hazard, a risk averse agent, and transfers. He argues that the optimal contract is maximally forgiving: it punishes only after the worst possible signal. In our model, monitoring more tasks to improve the efficiency of punishment (e.g., having a more informative worst signal) comes at the expense of productive activity. Moreover, an agent can prevent the worst outcome by performing only some of his tasks. That is, our monitoring structure has a moving support. Furthermore, in Mirrlees' model punishment essentially never occurs: the worst possible signal happens with negligible probability when the near-efficient action is taken. In our model, an agent is indeed forgiven unless his observed failures exceed some threshold, but he must be punished with positive probability if he is to exert any effort. Our results may be seen as occupying a middle ground between Mirrlees' worst-outcome contracts and making each agent the residual claimant of profits.

1.2 Bounded rationality interpretation

Viewed from a bounded rationality perspective, this paper departs from the common assumption in contract theory, and the economic literature at large, that an agent's memory has unlimited capacity and perfect recall.[3] Our model of team production among players with imperfect memories applies to tasks that are sufficiently difficult to describe that only a few of them can be stored in memory. A task in this view contains detailed information, such as a decision tree, that is necessary to complete it properly.[4] In such a case, it may be impossible to fully specify the details of the task in a convenient written form, and they must be enforced informally in equilibrium. When we say that a player "forgets" a task she had stored in memory, she actually forgets relevant details and is unable to complete the task properly. Even if she remembers the details, by ignoring them she can "botch" the task at no cost to herself. Another player can discover that she has botched the task only if he himself remembers these details.[5]

The literature in cognitive psychology has established that individual memories are imperfect, and, most importantly for models of interaction, that the *collective memory* of a group has very different properties than individual memory.[6] In particular, collective memory can be generated and maintained by collaborative recall processes such as *cross-cueing*, by which one individual's recall triggers a forgotten memory in another (Weldon and Bellinger 1997). The collective memory of the team serves here as a costly monitoring device. To contract upon an event in our setting it must be common knowledge. To this end, we assume that cross-cueing generates common knowledge; one justification for this is based on an underlying conceptual model that separates working memory, which is tightly bounded, from long-term memory, which is effectively unbounded. (Baddeley 2003 reviews the relevant psychological and neurological literature.) Information held in working memory (including cues to retrieve information from long-term memory) can be acted on, while information held in long-term memory can be used to verify claims about the past. In this interpretation, a player who has forgotten one of his promises from his working memory still holds it in his long-term memory. If another player holds his promise in her working memory, she can cross-cue him, reminding him of his promise to restore common knowledge.[7]

[3]Notable exceptions, typically in the decision-theoretic literature, include Dow (1991), Piccione and Rubinstein (1997), Hirshleifer and Welch (2001), Mullainathan (2002), Benabou and Tirole (2002) and Wilson (2004). There is also a literature on repeated games with finite automata which can be interpreted in terms of memory constraints (e.g., Piccione and Rubinstein 1993, Cole and Kocherlakota 2005, Compte and Postlewaite 2008, Romero 2010), as well as work on self-delusion in groups (e.g., Benabou 2008).

[4]Al-Najjar, Anderlini and Felli (2006) characterize finite contracts regarding "undescribable" events, which can be fully understood only using countably infinite statements. In this interpretation, to carry out an undescribable task properly, a player must memorize and recall an infinite statement. The related literature considers contracts with bounded rationality concerns relating to complexity—such as limitations on thinking through or foreseeing contingencies (e.g., Bolton and Faure-Grimaud Forthcoming, Maskin and Tirole 1999, Tirole 2009), communication complexity (e.g., Segal 1999), and contractual complexity (e.g., Anderlini and Felli 1998, Battigalli and Maggi 2002).

[5]We assume that the benefit of a task is in expectation, and that players cannot contract on their ex-post payoffs.

[6]A seminal paper by Miller (1956) suggests that the capacity of working memory is approximately 7 ± 2 "chunks." A chunk is a set of strongly associated information—e.g., information about a task. More recently, Cowan (2000) suggests a grimmer view of 4 ± 1 chunks for more complex chunks.

[7]Ericsson and Kintsch (1995) note, "the primary bottleneck for retrieval from LTM [long-term memory] is the scarcity of retrieval cues that are related by association to the desired item, stored in LTM." Here the review stage of the game provides the necessary retrieval cues. Smith (2003) shows that intending to perform a task later requires using working memory to monitor for a cue that the time or situation for performing the task has arrived.

Our model bears relations to literatures in cognitive psychology and organizational behavior. Remembering a promise (i.e., remembering one's intention to complete a task at a later point) is termed *prospective memory* in the theory of cognitive psychology; Dismukes and Nowinski (2007) study prospective memory lapses in the airline industry, noting that they are "particularly striking" because that industry has "erected elaborate safeguards... including written standard operating procedures, checklists, and requirements... to cross check each other's actions." In view of such difficulties, various theories of how to optimally store, recall, and share information have been proposed in the literature on organizational behavior; for example, consider Mohammed and Dumville (2001), Xiao, Moss, Mackenzie, Seagull and Faraj (2002) and Haseman, Nazareth and Paul (2005), which draw on the seminal work of Wegner (1987). Wegner develops the notion of *transactive knowledge*, the idea that while we cannot remember everything, we know who remembers what we need to know. That is, "memory is a social phenomenon, and individuals in continuing relationships often utilize each other as external memory aids to supplement their own limited and unreliable memories" (Mohammed and Dumville 2001). In our model, players know who is responsible for each task as well as who is responsible for monitoring the promiser. This bears a formal relationship to *transactive responsibility*, a concept that Xiao et al. (2002) introduce to study the division of responsibilities and cross-monitoring by trauma teams in hospitals.

2 The model

Consider first a loose overview of the model. Before the game starts, a contract is in place that governs the punishment each player will receive as a function of the messages sent at the end of the game. There are three stages:

1. *Promise-making.* Each player promises to complete certain tasks, and then memorizes some subset of the team's promises. Promises are public, but memorization is private.
2. *Task-completion.* Each memory slot fails with some probability, independently across slots. A player chooses some subset of her recalled tasks to complete. The other tasks she promised are botched.
3. *Review.* Completion of a task is unverifiable, but a task can (verifiably) be discovered to have been botched if it is monitored. Players monitor those tasks that they remember, and send a public report about their monitoring results. Based on these reports, each player is punished according to the contract.

For most of the paper we focus on the case of a two-player team, $\mathcal{I} = \{1, 2\}$. In Section 7, we extend the analysis to larger teams. A countably infinite set of *tasks* $\mathcal{X}$ is available to the team. Each task can be completed by one team member, who must memorize and recall detailed information about the task to complete it. Each player i has a bounded memory with M_i slots, each of which may be used to store a *promise* $(x, j) \in \mathcal{X} \times \mathcal{I}$ encoding a task x and the player j who promises to complete it. A single promise can be stored in at most one memory slot, so a player's memory state is an element of $\mathcal{M}_i = \{m_i \subset (\mathcal{X} \times \mathcal{I}) \big| |m_i| \leq M_i\}$. A player reaps a benefit b from each task that is properly completed by the team, but incurs cost c for each task he completes himself. Completing any given task is efficient but a player would rather not do it; i.e., $b < c < 2b$.

With a contract in place at the outset of the game (we formalize contracts in Section 2.2, below), the players enter the promise-making stage. Each player i publicly announces promises $\pi_i \subset \mathcal{X} \times \{i\}$. Given the collection of all promises, $\pi = \bigcup_j \pi_j$, each player privately decides which of these promises to memorize. Player i's memorization strategy is $\mu_i : 2^{\mathcal{X} \times \mathcal{I}} \to \Delta\mathcal{M}_i$. We assume that players cannot delude themselves; i.e., the support of $\mu_i(\pi)$ must be contained within π.[8]

By the task-completion stage, each promise that player i had memorized is recalled with probability $\lambda_i \in [0, 1]$, independently across promises. Her resulting memory state is $m_i \in \mathcal{M}_i$. A player cannot fulfill a promise for which she has forgotten the necessary details. Consequently, player i's decision strategy $d_i : \mathcal{M}_i \to \Delta 2^{\mathcal{X}}$ for which promises to fulfill can put positive probability only on promises contained in m_i.

At the review stage, each player monitors the tasks she remembers that her teammate promised (i.e., player i monitors promises $m_i \cap \pi_{-i}$), and publicly reports which of those promises went unfulfilled. Let $A_i \subset \mathcal{X} \times \{i\}$ be the set of promises that player i fulfilled, and let $\hat{m}_i \subseteq m_i \cap \pi_{-i} \setminus A_{-i}$ be the set of her teammate's unfulfilled promises that she reports. The collective memory, $\bigcup_j \hat{m}_j$, then contains the union of all reported unfulfilled promises. We assume that failure to complete a task is verifiable, and that only verified reports are incorporated into the collective memory.[9]

2.1 Equilibrium refinement: simple monitoring strategies

To determine whether she would like to fulfill some subset of her recalled promises, a player must be able to compute—at the task-completion stage—the conditional distribution over which subsets of her recalled promises will be monitored. We focus on a class of monitoring strategies that are a straightforward generalization of pure strategies, where any randomization (if necessary) is uniform.

Player i's monitoring strategy μ_i is *simple* if (i) her allocation of memory between own promises and monitoring is deterministic and (ii) she randomizes uniformly which promises to monitor, if the space allocated for monitoring is smaller than the number of promises made.

[8] Hence the memory process differs from Benabou (2008), which is interested in distortions of reality.

[9] This is in line with the literature on *cross-cueing* (e.g., Weldon and Bellinger 1997): a player triggers the memory of his teammate when he reports on the details of a task. Note that a player can be reminded of missing details to be convinced that a task has not been properly completed, but cannot be fully convinced by another player that *all* details of a task have been properly completed. That is, proper completion of a task cannot be made common knowledge, and is therefore not contractible.

Under simple monitoring strategies, player i's task-completion strategy need depend only on the number of promises she recalls, the contract, how many promises she made, and how many of those are being monitored. We assume she recalls these bare outlines of the promise-making stage perfectly, even if she cannot recall the promises made in greater detail.[10] This simplification avoids forcing players to memorize potentially complicated monitoring strategies over subsets of π_i in a setting in which they have bounded memory and imperfect recall; and as such, can be viewed as satisfying a technological constraint of memory. (In the classical interpretation, the assumption of simple monitoring strategies extends the setting to which our analysis applies.[11])

In an equilibrium in simple monitoring strategies, no player can profit by deviating to a non-simple monitoring strategy. Therefore, our focus on simple monitoring strategies serves as an equilibrium refinement and not a restriction on the set of strategies available.

2.2 Counting contracts

A contract, fixed at the outset of the game, determines a vector of punishments that will be applied at the end of the game. First, the contract can enforce any number of equilibrium promises using the threat of harsh punishments.[12] Second, if nobody deviated in the promise-making stage, then the contract yields a vector of punishments as a function of the collective memory at the end of the review stage, $V : 2^{\mathcal{X}\times\mathcal{I}} \to \mathbb{R}_-^{|\mathcal{I}|}$. The ex-post payoff of player i is

$$U_i = b\sum_j |A_j| - c|A_i| + V_i\Big(\bigcup_j \hat{m}_j\Big). \tag{1}$$

We study symmetric *counting contracts*, a straightforward and intuitive class of contracts, in which each player's punishment depends only on the number of her unfulfilled promises that are reported by her teammate. She can compute the distribution of this number using only the number of promises she recalls, how many promises she made, and how many of those are being monitored. Hence a counting contract is compatible with simple monitoring strategies.

Assumption 1 (Counting contracts). *Let $f_i \equiv |\bigcup_j \hat{m}_j \cap (\mathcal{X}\times\{i\})|$ denote the number of player i's unfulfilled promises that have been discovered. A contract must be a* counting contract *of the form $V_i(\bigcup_j \hat{m}_j) = v_i(f_i)$, where $v_i : \mathbb{I}_+ \to \mathbb{R}_-$.*

[10] This is one possible formalization of the sentiment in Wegner (1987) that "we have all had the experience of feeling we had encoded something... but found it impossible to retrieve."

[11] For example, when the constrained resource is time, blocks of time may be available in an i.i.d. fashion, but the agent can choose from among all his promises which tasks to complete. Though conceptually distinct from the bounded rationality interpretation in that it separates the task name from the opportunity to perform it, with simple monitoring strategies the analysis is identical.

[12] Alternatively, any number of promises can be part of a perfect Bayesian equilibrium under the following deviation response: if anyone promises a deviant set of tasks, nobody commits any promises to memory, yielding zero payoffs. Since players are indifferent to monitoring or not, this off-equilibrium play is sequentially rational.

Since a counting contract cannot punish a player for her report (which is verifiable), it follows that she is willing to fully disclose what she recalls of her teammate's promises. Similarly, with counting contracts and simple monitoring strategies, there is no benefit to making more promises than one intends to memorize. Without loss of generality, we focus on equilibria in which players fully memorize their own promises and fully disclose their monitoring observations. Henceforth we abuse terminology, using "contract" to refer to a counting contract, with or without its corresponding full-memorization, full-disclosure, simple monitoring perfect Bayesian equilibrium. A contract is *optimal* if it is Pareto optimal in the space of all such contracts satisfying these conditions.

3 The design problem

We now develop the problem of designing an optimal symmetric contract. In a symmetric contract, let p be the number of promises each player makes, and let F be the number of slots each player devotes to monitoring. Remember that under simple monitoring strategies, player $-i$'s monitoring is uniform, and so conditional on being asked to fulfill x of her promises, player i is indifferent over which x promises to fulfill. Therefore, we may represent a task-completion strategy using $s : \{0, \dots, p\} \to \{0, \dots, p\}$, where $s(k)$ maps the number of her promises that she recalls, k, to the number of tasks she performs, $s(k)$.[13] Naturally, the strategy must satisfy $s(k) \leq k$.

To determine whether a strategy s is incentive compatible, we must know the probability distribution over the number of discovered unfulfilled promises f conditional on $s(k)$ for each $k = 0, \dots, p$. Given F and p, if a player fulfills a of her promises, the probability that her teammate will find f of her unfulfilled promises is given by the compound hypergeometric-binomial distribution

$$g(f, a) = \sum_{k=f}^{F} \frac{\binom{p-a}{k}\binom{a}{F-k}}{\binom{p}{F}} \binom{k}{f} \lambda^f (1-\lambda)^{k-f}. \tag{2}$$

To interpret Eq. 2, observe that in order to discover f unfulfilled promises of player i, player $-i$ must have drawn $k \geq f$ promises from the $p-a$ promises player i failed to fulfill, and $F-k$ promises from the a promises player i fulfilled; this is described by a hypergeometric distribution. Of these k promises, player $-i$ must then recall exactly f; this is described by a binomial distribution. This distribution is studied by Johnson and Kotz (1985) and shown by Stefanski (1992) to have a monotone likelihood ratio property: $g(f, a)/g(f, a-1) < g(f-1, a)/g(f-1, a-1)$ for all a, f. Hence an increase in the number of tasks completed yields a first-order stochastic improvement in the number of unfulfilled promises discovered.

[13] That is, s expresses the task completion strategy d_i in a simpler form. By restricting $s(k)$ to be a number rather than a random variable, we use the fact that to randomize, the player must be indifferent, but then it would be optimal for her to put probability 1 on the highest number of tasks in the support of her randomization.

The incentive constraints for implementing a task-completion strategy s are

$$\sum_{f=0}^{F} v(f)\Big(g\big(f, s(k)\big) - g(f, \ell)\Big) \geq \big(s(k) - \ell\big)(c - b) \quad \text{for all } \ell \leq k, \text{ and all } k. \tag{3}$$

We call these "downward" constraints when $\ell < s(k)$, and "upward" constraints when $s(k) < \ell \leq k$. An optimal contract v, combined with the optimal number of promises p, monitoring slots F, and task-completion strategy s, maximizes expected social benefits net of punishments, subject to incentive compatibility:

$$\begin{aligned} \max_{v,p,F,s} \sum_{a=0}^{p} \binom{p}{a} \lambda^a (1-\lambda)^{p-a} \Big(s(a)(2b - c) + \sum_{f=0}^{F} v(f) g(f, s(a))\Big) \\ \text{s.t. } v(f) \leq 0 \text{ for all } f, \text{ and Eq. 3.} \end{aligned} \tag{4}$$

In particular, optimality requires that the contract v implements the strategy s at minimum punishment cost, subject to the incentive constraints. Two insights may be gleaned directly from this design problem:

1. *If a positive number of promises should be completed in equilibrium, then the upward incentive constraints must be slack at the optimal contract.* Indeed, if a player is indifferent between performing $s(k)$ tasks and some larger number of tasks ℓ (where $s(k) < \ell \leq k$), then it would be both incentive compatible and socially beneficial for her to complete ℓ tasks.

2. *An optimal task-completion strategy s is an increasing step function, and calls for completing as many promises as are remembered whenever it jumps.* Formally, $k \geq \ell$ implies $s(k) \geq s(\ell)$; and $s(k) > s(k-1)$ implies $s(k) = k$. This follows from a simple revealed preference argument: if doing $s(\ell)$ is preferred to doing any $\ell' \leq \ell$ tasks when ℓ tasks are remembered, then $s(\ell)$ remains preferred to any $\ell' \leq \ell$ tasks when $k \geq \ell$ tasks are remembered

4 Benchmark: linear contracts

We begin by studying the benchmark case of symmetric *linear contracts*, in which a player is punished a constant amount for each unfulfilled promise discovered by the team. We suppose in this section that there is a per-task punishment bound of $\underline{v} < 0$; that is, we look at contracts of the form $v_i(f_i) = \tilde{v} f_i$, where $\tilde{v} \in [\underline{v}, 0]$. While very simple, such contracts may be natural in settings where the punishments are imposed in a linear fashion by third parties. For instance, doctors on a medical team may be ethically bound to report each others' errors to the affected patients, each of whom then reacts with an i.i.d. probability of suing for malpractice damages.

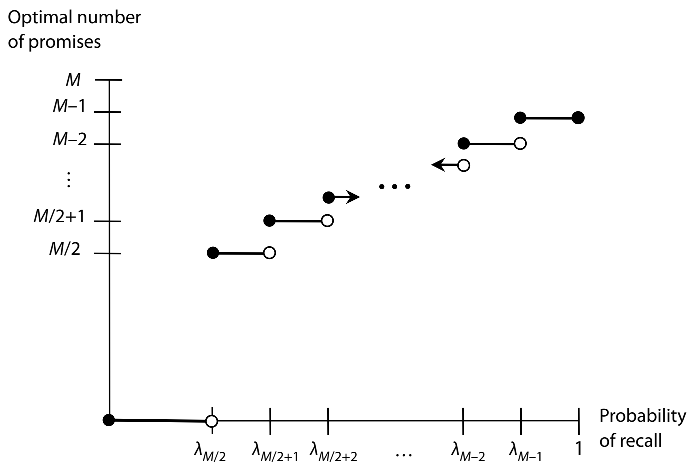


FIGURE 1: OPTIMAL LINEAR CONTRACT REGIMES. Here, $\lambda_{M/2} = \max\{\frac{c-b}{b}, \frac{b-c}{\underline{v}}\}$. All λ-ranges shown are nonempty if $-b \leq \underline{v} \leq (M-1)(b-c)$.

The main theorem of this section characterizes optimal symmetric linear contracts for M even.[14]

Theorem 1 (Linear contracts)**.** *Suppose M is even. If $\lambda > \max\{\frac{c-b}{b}, \frac{b-c}{\underline{v}}\}$, then the optimal number of promises, p, and the optimal per-task punishment $\tilde{v}$, are given by*[15]

$$p = \left\lfloor \frac{M}{1 + \frac{b-c}{\lambda \underline{v}}} \right\rfloor \text{ and } \tilde{v} = \frac{p(b-c)}{\lambda(M-p)}. \tag{5}$$

Each player completes as many of her promises as she recalls, giving expected social welfare $2p(b-c+b\lambda)$. If $\lambda < \max\{\frac{c-b}{b}, \frac{b-c}{\underline{v}}\}$, then players optimally make no promises.

Under the optimal linear contract, each player fulfills as many promises as she remembers, so we say the contract satisfies "promise keeping." Because a linear contract treats each task separately and symmetrically, if she is willing to fulfill any single promise, then she is willing to fulfill all her promises. Thus she effectively faces a single incentive constraint at the task-completion stage.[16] In an optimal contract, this incentive constraint must bind, since otherwise the punishments could be made less severe without violating incentive compatibility. Whenever $\lambda > \frac{c-b}{b}$, p should be maximized subject to the constraints. If $\lambda < \frac{b-c}{\underline{v}}$, the punishment bound is too restrictive to incentivize completing any tasks. When the bound $\underline{v}$ is very restrictive, the optimal number of promises may be smaller than $M-1$ even if $\lambda = 1$.

[14]There may be superior asymmetric linear contracts, but they will not differ from the optimal symmetric contract by more than a task per player. Similarly, for M odd all optimal linear contracts, symmetric or otherwise, will be close to the optimal symmetric linear contracts for $M-1$ and $M+1$. See footnote 19, below. Note that here truthful reporting is a restriction: otherwise, by underreporting, a nonlinear contract can be emulated.

[15]Define $\lfloor\cdot\rfloor$ as the "floor" function $\lfloor y \rfloor \equiv \max\{\hat{y} \in \mathbb{I} : \hat{y} \leq y\}$.

[16]The IC simplifies to $b-c \geq \lambda \min\{\frac{M-p}{p}, 1\} v$, where $\min\{\frac{M-p}{p}, 1\}$ is the marginal probability a task is monitored.

How the optimal number of promises changes with the probability of recall is illustrated in Figure 1. The optimal number of promises decreases in $\frac{b-c}{\lambda \underline{v}}$, the ratio of net loss from completing a task to the expected punishment for not completing it when the worst case punishment is used. Hence the optimal number of promises is increasing in λ, *ceteris paribus.* When λ is very low, the players optimally make no promises to avoid nearly inevitable punishments. As λ rises, it reaches a threshold at which it is both feasible and optimal to make some promises. At this threshold, in order to avoid leaving memory slots unused, the players devote at least half of their memory slots to promises. Therefore, either $p = 0$ or $p \geq \frac{1}{2}M$. As λ rises further, the amount of memory devoted to monitoring decreases—and the optimal number of promises increases. Note that $p \leq M - 1$ in any equilibrium, in order to leave at least one slot for monitoring to maintain incentive compatibility.

The key features of linear contracts are that players optimally make as many promises as they can incentivize given the bound on punishments, and that they fulfill as many promises as they recall. In the following section, we consider nonlinear counting contracts, and show that even if punishments may be unboundedly severe, it will not always be optimal make the maximum number of promises, or even to fulfill all the promises that are recalled.

5 General counting contracts

The linearity assumption made in the previous section simplified analysis, since a linear contract treats each task separately. However, under a linear contract there is a significant likelihood that the players will not recall all of their promises, which means they face a significant likelihood of being punished. Intuitively, a linear contract might be improved on by "forgiving" a player who completes all but the last few of her promised tasks. Of course, she will not fulfill any promises for which she will be forgiven, so some of her promises will be "empty." The drawback of such a forgiving contract is that, in the unlikely event in which she recalls all of her promises, she will not fulfill all of them. The benefit is that in the very likely event that she does not recall all of her promises, she will not be punished too severely.

In this section we analyze non-separable contracts, without any bound on the severity of punishments. In a non-separable contract, a player's punishment can depend in an arbitrary way on the number of her unfulfilled promises that are recalled by her teammate. The main tradeoff in designing optimal non-separable contracts is between using information efficiently and ensuring that a player recalls sufficiently many promises. To provide incentives for a player to complete any given number of recalled tasks, it is most cost-effective to use the most informative signal for punishment. Mirrlees (1974, 1999) proposed this basic intuition, but our model raises the complication that a player may be able to move the support of the monitoring distribution by fulfilling enough promises. If a player recalls a small number of promises, then being punished only for the worst outcome (the maximal number of unfulfilled promises are discovered) provides the most efficient incentives. However, if a player happens to recall a large number of promises, she may have incentive to fulfill only enough of them that the worst outcome cannot arise. Thus she may leave some promises unfulfilled; these are *empty promises*. A memory slot devoted to an empty promise is a memory aid: it helps the player recall more promises, yielding a first-order stochastic improvement in the number of promises she recalls. At the same time, an empty promise uses up a memory slot that could be used towards obtaining a more informative monitoring signal. The better the players' memories, the more slots they devote to "earnest promises" and the fewer slots they need devote to monitoring and empty promises.

In Section 5.1 and Section 5.2, we develop the following properties of optimal contracts. To preview these results, let p^* be the maximum number of promises that a player will ever actually fulfill. Properties 1–4 below hold for all memory sizes. Properties 5–7 are shown for $M \leq 5$, as well as for *maximally forgiving* contracts under all memory sizes. Properties 5–6 are shown for all memory sizes when contracts are constrained to be *decreasing convex* (deliver increasingly worse punishments for larger numbers of unfulfilled promises). In Section 5.3 we point out that our results are robust to two key generalizations of the model.

1. If λ is sufficiently high, then the optimal contract is linear with $p^* = p = M - 1$ and $F = 1$;
2. If λ is sufficiently low, then it is optimal to do nothing;
3. For a range of parameters, it is optimal for players to make empty promises;
4. Players make empty promises (with $p^* < p$) if and only if they make less than the maximum number of promises ($M - 1$);
5. Each player performs as many tasks as she recalls up to a cutoff p^*;
6. The optimal contract forgives up to a threshold, and punishments increase linearly thereafter;
7. The optimal number of promises (p) and the promise-completion cutoff (p^*) increase in λ, while empty promises ($p - p^*$) decrease in λ; and social welfare is increasing in λ.

5.1 Empty promises

We say that a strategy is *promise-keeping* if $s(a) = a$ for all $a \leq p$, and has *empty promises* otherwise. The following lemma shows that promise keeping is optimally implemented by a linear contract.

Theorem 2 (Promise-keeping and linear contracts)**.** *For any M and any p, promise keeping is optimally implemented by a linear contract. Moreover, when λ is sufficiently high or low, a promise-keeping linear contract is socially optimal:*

(i) There exists $\overline{\lambda} < 1$ such that for all $\lambda \geq \overline{\lambda}$, the unique optimal contract is linear with $v(f) = \frac{M-1}{\lambda}(b-c)f$, and promise-keeping with the maximal number of promises ($p = M-1$).

(ii) There exists $\underline{\lambda} > 0$ such that for all $\lambda \leq \underline{\lambda}$, there is an optimal contract that is linear with $v(f) = 0$ for all f, and degenerately promise-keeping with zero promises.

To see why promise keeping is optimally implemented by a linear contract, note that all the downward incentive constraints can be made to bind under a linear contract, thereby minimizing punishments. Intuitively, when λ is very high the players expect to recall almost all of their promises. Because they don't expect to incur punishments too often, it is optimal to maximize the number of promises made by setting $p = M-1$. But with only one monitoring slot, every task is treated identically, so the contract is linear (and takes the form detailed in Theorem 1). When λ is very low, players expect to recall few or none of their promises. Rather than risk incurring punishments, it is better not to do any tasks at all.

Between these extremes, however, it may be optimal for players to make empty promises, even in the region where promise-keeping gives positive social utility (i.e., when $\lambda \geq \frac{c-b}{b}$). To state the next theorem, let $p^* \equiv \max_a s(a)$ be the largest number of promises that are ever fulfilled under strategy s.

Theorem 3 (Empty promises)**.** *For an intermediate range of λ's, the optimal contract is nonlinear and implements empty promises. In particular, for any memory size M,*

(i) There exists $\alpha_M \in (1,2)$ such that if $c < b\alpha_M$, there exists $\tilde{\lambda} > \frac{c-b}{b}$ so that for all $\lambda \in (\frac{c-b}{b}, \tilde{\lambda})$, making and keeping the maximal number of promises $(M-1)$ yields positive social utility, but is dominated by making roughly half as many promises ($\lfloor \frac{M+1}{2} \rfloor$) and fulfilling just one whenever any promises are remembered.

(ii) If the optimal contract implements fulfilling a positive number of promises, players make empty promises (with $p^* < p$) if and only if $p < M-1$; that is, players "under-deliver" on their promises if and only if they "under-promise."

Hence there is generally a gap between the highest number of promises a player is willing to complete, and the number of promises that she actually made. Moreover, players should make as many promises as possible if and only if they intend to follow through on them. By memorizing more promises than she plans to fulfill, a player attains a first order stochastic improvement in the number of tasks she will complete according to her plan. However, the corresponding increase in the number of promises she leaves unfulfilled will lead her to expect a more severe punishment unless the contract is *forgiving*: if it does not punish her when the other players find only a "small" number of her unfulfilled promises. A forgiving contract therefore allows players to use empty promises as memory aids, and minimizes punishment in the very likely event that players cannot recall all of

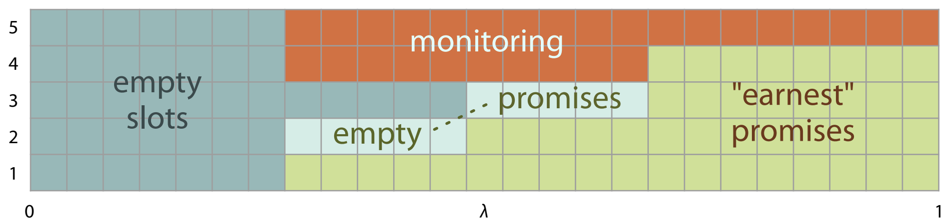


FIGURE 2: OPTIMAL MEMORY ALLOCATION FOR $M = 5$. Each column represents the optimal memory allocation for a particular value of λ, in steps of 0.04 from 0.02 to 0.98, where $b = 2$ and $c = 3$. In row a, $s(a) = a$ in the "earnest" promises region, while $s(a) < a$ in the empty promises region. Any cutoff strategy can be represented in this format.

their promises. Of course, if a player makes the maximal number of promises ($p = M - 1$), then the contract cannot be forgiving, since it must punish her when it finds even one unfulfilled promise. While finding a large number of unfulfilled promises is an informative signal of moral hazard, it may not arise with positive probability if a player completes all but a few of her promised tasks. Promise keeping is thus costly to implement; and depending on the quality of memory, a contract that enforces keeping all of one's promises can be dominated socially by a forgiving contract under which players keep just a single remembered promise.

Such a task-completion strategy has the form of a *cutoff* strategy: we call s a cutoff strategy if $s(a) = a$ for $a \leq p^*$ and $s(a) = p^*$ for all $a > p^*$. We show below that when $M \leq 5$, the optimal contract always implements cutoff strategies, that both the number of promises and the cutoff increase in λ, and that both monitoring and the number of empty promises decrease in λ. A specific example illustrating these comparative statics is visualized in Figure 2. Moreover, the optimal contract leaves approximately the same number of monitoring slots as empty promises, and is *maximally forgiving*: it punishes a player only when the maximal number of unfulfilled promises has been discovered (F).

Theorem 4. *Suppose $M \leq 5$. Then,*

(i) For any λ, the optimal contract implements cutoff strategies;

(ii) The optimal contract is maximally forgiving, with $p - p^* \leq F \leq p - p^* + 1$;

(iii) Both p and p^* increase in λ, while both F and $p - p^*$ decrease in λ;

(iv) Social welfare under the optimal contract is increasing (strictly whenever $p^* > 0$) and piecewise concave in λ.

When $M \leq 5$, it can be shown using a duality argument that for any combination of p, p^*, F, and λ, the optimal contract is maximally forgiving and implements a cutoff strategy. Moreover, we show that the resulting social value of a maximally forgiving contract satisfies single-crossing and concavity properties for any M. This is illustrated for a specific example in Figure 3, which graphs the conditionally optimal social welfare for each fixed $p^* \in \{1, 2, 3, 4\}$.

In general, it is difficult to analytically characterize the optimal punishment schedule without restrictions on the contract space. When $M > 5$, we have numerical examples where the optimal strategy is not a cutoff, or the optimal contract is non-monotone, for particular combinations of b, c and λ. More generally, in a principal-agent setting with transfers and where there is neither private information nor capacity constraints, Grossman and Hart (1983) suggest that without a strong assumption on the distribution of outputs (which is not satisfied by the distribution g), one may only conclude that the optimal wage schedule is not everywhere decreasing. However, it may be natural to restrict the contract space *a priori*, as in the next subsection.

5.2 Convex contracts

Contracts which deliver increasingly large punishments for larger numbers of unfulfilled promises may be a focal class to consider. Such *decreasing convex (DC)* contracts satisfy the restriction $v(f) - v(f+1) \geq v(f-1) - v(f) \geq 0$. Convex contracts may be natural in settings where punishments are imposed by third parties who are more inclined to exact punishment if they perceive a consistent pattern of failures. Conversely, a non-convex contract may be particularly difficult to enforce via an affected third party, since it would require leniency on the margin for relatively large injuries.

For arbitrary memory size M, we show that DC contracts optimally induce task-completion strategies that have a cutoff form. Furthermore, the optimal such contract forgives empty promises up to some failure threshold, and increases punishments linearly thereafter. For the special case of maximally forgiving contracts (as in Theorem 4), players make more total promises and fewer empty promises as λ increases. Figure 4 illustrates these comparative statics for $M = 7$ and $M = 9$.

Theorem 5. *For any M and any λ, every DC contract induces a cutoff strategy. Moreover,*

(i) An optimal DC contract uses a kinked-linear punishment schedule, with $v(f) = 0$ for all f smaller than some failure threshold $\hat{f} \leq F$.

(ii) Among maximally forgiving contracts ($\hat{f} = F$), optimally both p and p^* increase in λ, and $p - p^*$ decreases in λ. The social welfare of the optimal contract is increasing and piecewise concave in λ.

To prove this, we show that properties of the failure-detection distribution g imply that if the contract is convex in the number of unfulfilled promises that player $-i$ discovers, then player i's conditional expected punishment is convex in the number of her promises she fails to fulfill. This implies that if player i prefers completing $\tilde{p}$ promises over $\tilde{p}-1$ promises, then she must also prefer completing $\tilde{p}-k$ promises over $\tilde{p}-(k+1)$ promises for all $k=1,\ldots,\tilde{p}-1$. We then use a duality argument to show that the Lagrange multipliers for the convexity constraints in the contracting problem imply a recursion that can be used to solve for the optimal expected punishment. That expression can be written in terms of the expected number of discovered unfulfilled promises that exceed a threshold for punishment, and is implementable by a kinked-linear punishment schedule.

5.3 Robustness

Monitoring technology For simplicity, we have assumed that the probabilities of recalling one's own promise or a teammate's promise are the same. More generally, the performance and monitoring technologies may differ. Our qualitative results do not change if the probability of performing a task, λ^{p}, is allowed to differ from the probability of monitoring a task, λ^{m}, because the problem of implementing a strategy at minimal cost depends only on λ^{m}. Moreover, for the special case of maximally forgiving contracts in Theorem 2, Theorem 4, and Theorem 5(ii), the value of an optimal contract is entirely independent of the monitoring technology.[17] Intuitively, any reduction in λ^{m} can be perfectly compensated by increasing the punishment $v(F)$ in the right proportion, leaving the expected punishment and all incentives unchanged. The comparative statics there on earnest and empty promises are driven entirely by performance ability λ^{p}.

Strict incentives for monitoring In our setting, players are indifferent over whether to monitor each other. However, our results on the optimality of empty promises continue to hold even if players require strict incentives for monitoring. In Rahman (2009), to provide incentives for monitoring a player may occasionally shirk just to "test" the monitor. Our model generates optimal "shirking"—in the form of empty promises—with positive probability even in the absence of monitoring costs. These empty promises can also be repurposed to provide strict incentives to monitor. If, after players announce their monitoring observations, they are then asked to self-report any promises they recalled but failed to fulfill, it becomes possible to detect and punish failures of monitoring. For strict incentives, a player need be punished only infinitesimally for failing to monitor. This mechanism also ensures robustness to small costs of reporting or verifying failed tasks.

6 Asymmetric players

Previous sections assumed that both players had equally good memories, and studied contracts in which the players shared equally in the responsibilities both for accomplishing tasks and for monitoring. But if the players' memory abilities differ (along two dimensions: memory capacity and

[17] Observe that in Eq. 2, $g(F,a)=(\lambda^{\mathrm{m}})^{F}\binom{p-a}{F}/\binom{p}{F}$.

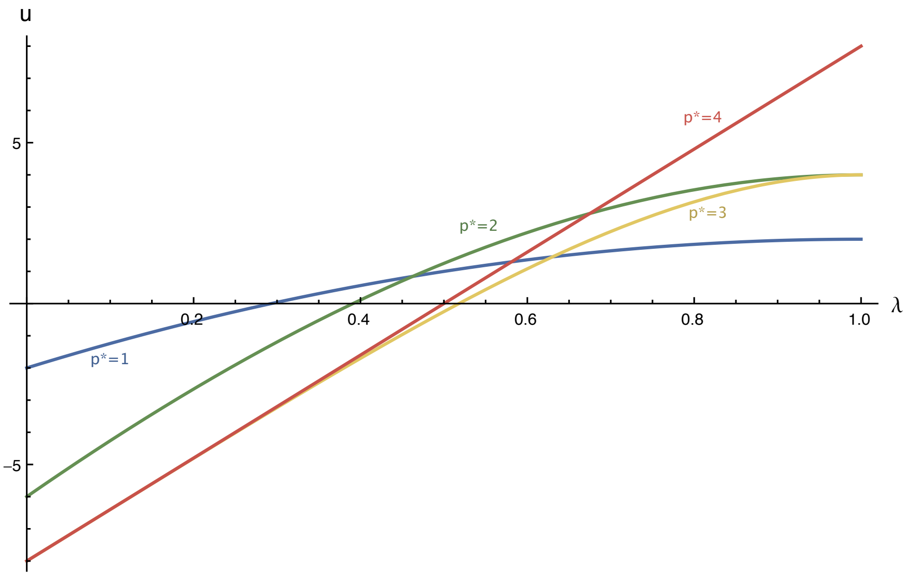


FIGURE 3: SOCIAL WELFARE ENVELOPE FOR $M = 5$. For each $p^* \in \{1, \ldots, M-1\}$, the value of the best strategy is plotted as a function of λ. In each case the best strategy is a cutoff strategy.

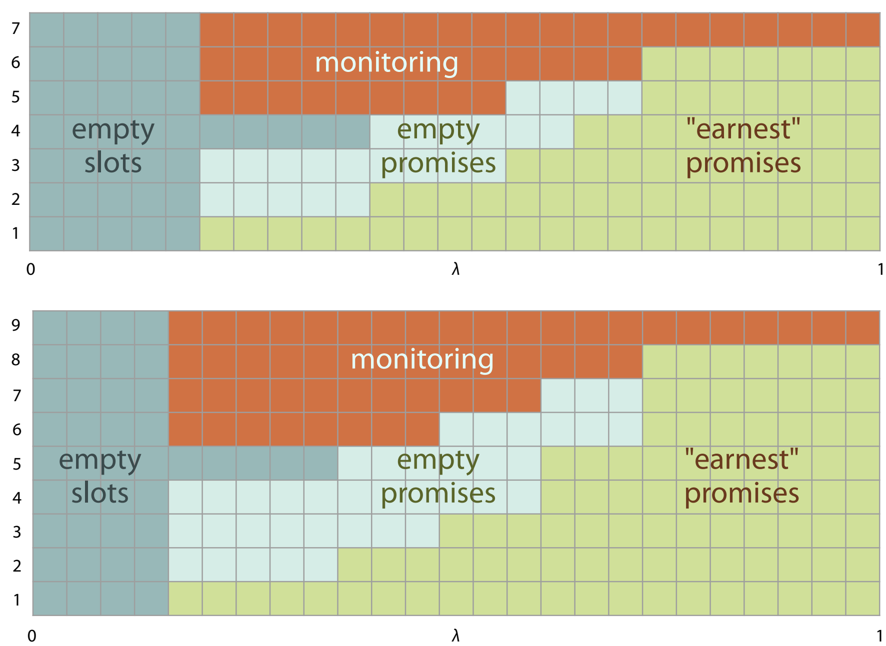


FIGURE 4: OPTIMAL MEMORY ALLOCATION UNDER MAXIMALLY FORGIVING CONTRACTS WITH $M = 7$ (TOP) AND $M = 9$ (BOTTOM). See Figure 2 for explanation. For these particular parameters, the maximally forgiving contracts are actually optimal among all contracts.

probability of recall), how should task-accomplishing and monitoring responsibilities be allocated? That is, which player should be given more "supervisory responsibility"?

In this section we endow each player i with an idiosyncratic memory capacity M_i and recall probability λ_i. To study the comparative statics of responsibility, we focus on linear contracts. We show that greater supervisory responsibility is optimally assigned to the less productive team member. Indeed, if the disparity in memory ability is large enough, it can be optimal for the weaker player to become a full-time supervisor who specializes entirely in monitoring, and pay a lump sum "salary" to the stronger player to accept responsibility for performing tasks. This accords with the "Dilbert Principle" of Adams (1996), which states "the most ineffective workers are systematically moved to the place where they can do the least damage: management."

To gain graphical intuition for this effect, we drop the integer restriction on the number of promises. Let us reintroduce a bound $\underline{v}$ on per-task punishments. Since the players randomize uniformly over which promises to monitor, the probability that player $-i$ monitors any given task of player i is $\lambda_{-i}\min\big\{\frac{M_{-i}-p_{-i}}{p_i},1\big\}$, where λ_{-i} is the recall probability of player $-i$. An optimal asymmetric linear contract solves

$$\begin{aligned}
&\max_{(v_i,p_i),\, i=1,2}\ \Big\{\sum_{i=1,2} p_i\big(\lambda_i(2b-c)+(1-\lambda_i)\min\big\{\tfrac{M_{-i}-p_{-i}}{p_i},1\big\}\lambda_{-i}v_i\big)\Big\}\\
&\text{subject to, for } i=1,2:\\
&\text{Feasibility:}\quad \underline{v}\le v_i\le 0 \text{ and } 0\le p_i\le M_i,\\
&\text{IC}_i:\quad b-c\ge \lambda_{-i}\min\big\{\tfrac{M_{-i}-p_{-i}}{p_i},1\big\}v_i \text{ if } p_i>0.
\end{aligned}\tag{6}$$

In the objective, each promise player i makes reaps a benefit of $2b-c$ for the team when it is remembered and an expected punishment of $\min\big\{\frac{M_{-i}-p_{-i}}{p_i},1\big\}\lambda_{-i}v_i$ when it is forgotten.

Note that for an optimal symmetric contract the option to not make promises implicitly guarantees individual rationality, but, with asymmetric contracts, maximizing social surplus no longer guarantees that both players earn non-negative expected utility. If the players can make ex ante transfer payments, however, then the solution to Eq. 6 is always a Pareto improvement over autarky, since a player who would earn negative expected utility can be compensated by a lump sum transfer.

In an environment with ex ante transferable utility, we find that each player's responsibility for accomplishing tasks is decreasing in his disadvantage relative to the other player in both memory capacity and recall probability. Hence the less able player is allocated relatively more "supervisory responsibility." If the disparity in memory is extreme, our result indicates that the less able player is endogenously selected as the full-time supervisor, in accordance with the Dilbert principle.

Theorem 6. *Suppose $\lambda_i\ge\max\{\frac{b-c}{\underline{v}},\frac{c-b}{b}\}$ and $\underline{v}\le\frac{b-c}{\lambda_i}\frac{b-c+\lambda_i b}{b-c+\lambda_{-i}b}$, for $i=1,2$. Then in the class of linear contracts with ex ante transfers, optimal contracts exhibit these comparative statics in both*

λ_i and M_i: p_i is increasing, F_i is decreasing, p_{-i} is decreasing, and F_{-i} is increasing (all strictly if $p_i > 0$ and $p_{-i} > 0$). In addition, if $\frac{b-c}{\lambda_{-i}\underline{v}} M_i \geq M_{-i}$ and $\lambda_i \geq \lambda_{-i}$, then $p_{-i} = 0 < p_i$.

Relative to the symmetric setting, the player with the worse memory benefits not only from the greater number of promises her teammate optimally makes, but also from a reduction in the number of promises she will make. This is because she must increase her monitoring in order to incentivize her more capable teammate to accomplish a greater number of tasks. Because of the capacity constraint on the total number of monitoring and productive resources, the less able player is best employed as a monitor of the more able player's tasks.

If instead no ex ante transfers are allowed, then an optimal contract should yield non-negative utility to both players. To account for this, we introduce the following individual rationality constraints for $i = 1, 2$:

$$\text{IR}_i: \quad p_i\lambda_i(b-c) + p_{-i}\lambda_{-i}b + (1-\lambda_i)\lambda_{-i}p_i \min\left\{\tfrac{M_{-i}-p_{-i}}{p_i}, 1\right\}v_i \geq 0. \tag{7}$$

An optimal asymmetric contract without ex ante transfers solves Eq. 6 subject to these additional constraints.

Theorem 7. *Suppose $\lambda_i \geq \max\{\frac{b-c}{\underline{v}}, \frac{c-b}{b}\}$ for $i = 1, 2$. Then in the class of linear contracts without ex ante transfers, for any $M > 0$ and $\lambda \in (0,1)$, there exists open neighborhood of (M, λ) such that the comparative statics of Theorem 6 hold if (M_i, λ_i) is in this neighborhood for $i = 1, 2$. Furthermore, in this neighborhood player i's expected utility is decreasing and player $-i$'s expected utility is increasing in M_i and λ_i.*[18]

We sketch the proof of both these results, which are detailed in the appendix. As in Section 4, the incentive compatibility constraint must bind for any player who makes at least one promise; otherwise, the punishment v_i could be lessened slightly, raising expected payoffs without violating any constraints. Substituting each binding IC into both the objective and IR, and incorporating the punishment bound into IC, yields a reduced form linear problem for optimal linear contracts without ex ante transfers:

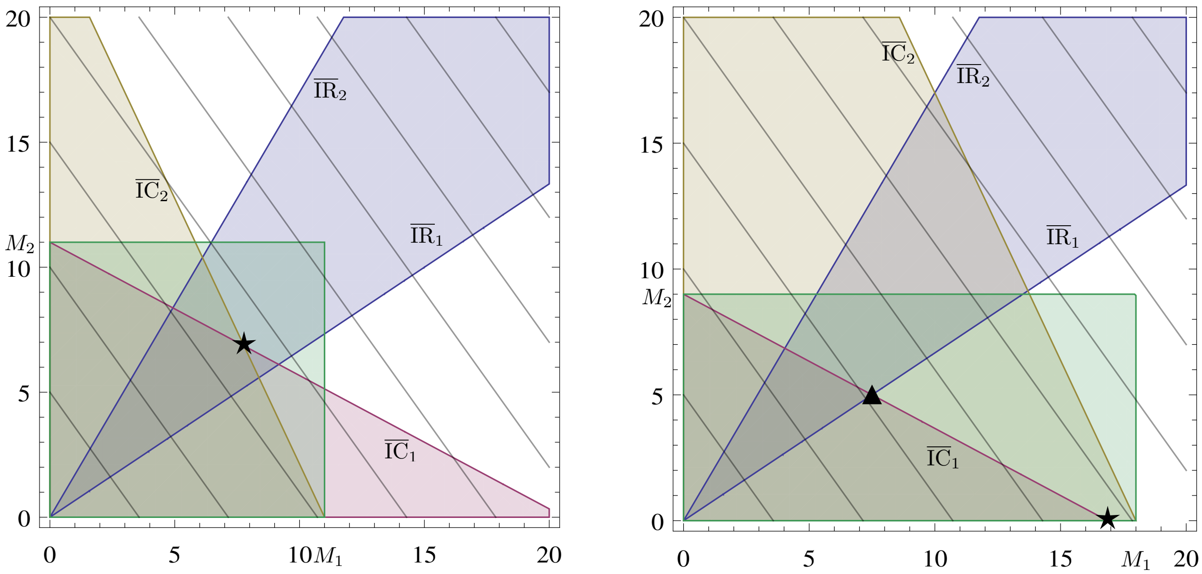
$$\max_{p_1, p_2} \left\{p_1(b-c+\lambda_1 b) + p_2(b-c+\lambda_2 b)\right\} \quad \text{s.t.} \tag{8}$$

$$\text{Feasibility: } 0 \leq p_i \leq M_i \text{ for each } i,$$

$$\overline{\text{IC}}_i: \quad b-c \geq \lambda_{-i}\underline{v}\min\left\{\frac{M_{-i}-p_{-i}}{p_i}, 1\right\} \text{ if } p_i > 0, \text{ and} \tag{9}$$

$$\overline{\text{IR}}_i: \quad p_i(b-c) + p_{-i}\lambda_{-i}b \geq 0.$$

[18] Formally, defining $\sigma_i \equiv \frac{b-c}{\lambda_i \underline{v}} \in (0,1)$ for $i = 1, 2$ and $\gamma \equiv -\frac{b}{\underline{v}} > 0$, the conclusion carries over for $\frac{M_1}{M_2} \in \left(\frac{\sigma_1\sigma_2+\gamma}{\sigma_1(1+\gamma)}, \frac{\sigma_2(1+\gamma)}{\sigma_1\sigma_2+\gamma}\right)$ and $\frac{\sigma_2}{\sigma_1}\frac{\sigma_1-\gamma}{\sigma_2-\gamma} \in (\sigma_2, \frac{1}{\sigma_1})$. See the formal proof in the appendix.



(a) An interior optimum

(b) An endogenous supervisor

Figure 5: Optimal asymmetric linear contracts. In both figures, the horizontal axis measures promises of player 1, and the vertical axis measures promises of player 2. The grey diagonal lines are social indifference curves. The optimal linear contract with ex ante transfers implements the promise vector ★ (ignoring integer problems). The parameters common to both figures are $b = 2$, $c = 3$, $\overline{v} = -2.5$, $\lambda_1 = .85$, and $\lambda_2 = .75$. In Figure 5(a), where $M_1 = M_2 = 11$, the optimal contract with ex ante transfers also satisfies individual rationality without transfers. In Figure 5(b), where $M_1 = 18$ and $M_2 = 9$, the optimal contract with ex ante transfers selects player 2 as an endogenous supervisor who devotes his entire memory to monitoring player 1. The the optimal linear contract without ex ante transfers, in contrast, implements the promise vector ▲.

The problem is visualized in Figure 5(a), which depicts the promises of player 1 on the horizontal axis and those of player 2 on the vertical axis. These are bounded by the rectangle corresponding to their respective memory capacities. Note that the set of non-zero promise pairs satisfying $\overline{\mathrm{IR}}_i$ is guaranteed to be nonempty if $\lambda_i \geq \frac{c-b}{b}$ for $i = 1, 2$. Moreover, it is possible to satisfy each $\overline{\mathrm{IC}}_i$ when player $-i$ monitors maximally with maximal punishments under the assumption that $\lambda_i \geq \frac{b-c}{\underline{v}}$. Under the assumptions of Theorem 7, the intersection of $\overline{\mathrm{IC}}_1$ and $\overline{\mathrm{IC}}_2$ occurs at a strictly positive promise pair, within both the $\overline{\mathrm{IR}}$ region and the memory bounds. As seen in Figure 5(a), the social indifference curve then selects the intersection of the incentive compatibility constraints to be the optimal promise pair. Since $\lambda_1 > \lambda_2$ and $M_1 = M_2$, player 1 has the superior memory and hence bears a larger task burden.

Without ex ante transfers, the individual rationality constraints require each player to complete sufficiently many tasks to make the optimal asymmetric contract palatable to her teammate. With ex ante transfers, however, the individual rationality constraints are relaxed, and if the incentive compatibility constraints do not intersect in the positive quadrant, one player can optimally become a full-time supervisor who pays the other player to accomplish tasks. An example is seen in Figure 5(b), where player 1's memory capacity is so large that player 2 optimally devotes his entire memory to monitoring.

7 Organizational implications

Thus far we have considered only 2-player teams. This section expands the analysis to larger teams, demonstrating that our results extend naturally to n-player teams. Motivated by applications to organizational structure, we propose two extended interpretations of our framework.

7.1 More than two players

Consider a team $\mathcal{I} = \{1, \dots, n\}$, $n > 2$. In this more general setting, we scale the benefits of each task so that each player reaps a benefit b from a task she performs herself but a benefit $\frac{b}{n-1}$ from a task performed by any other player. In any n-player symmetric equilibrium in which each task is monitored by a single player, it does not matter whose memory slot is used to monitor whom. Hence our results for two players extend naturally to a class of symmetric n-player equilibria, as formalized in Theorem 8, below.

We generalize simple monitoring strategies to a larger team as follows. A monitoring strategy is *simple and symmetric* if (i) each player's allocation of memory between own promises and monitoring is deterministic; (ii) there exists k such that the total number of slots allocated to monitoring each player is exactly k; (iii) each promise is monitored by at most one player; and (iv) the probability that any particular k promises of player i, $\{(x_1, i), \dots, (x_k, i)\} \subset \pi_i$, are monitored is identical. We say a contract is *optimal* if it is optimal in this class.

Theorem 8. *Consider any symmetric counting contract in simple monitoring strategies with $n = 2$, number of promises p, number of monitoring slots F, task completion strategy s, and punishment schedule v. Then for $n > 2$ there exists a contract in simple and symmetric monitoring strategies with the same p, F, s, and v for each player. The converse also holds. Moreover, the contract for $n = 2$ is optimal if and only if the contract for $n > 2$ is optimal.*

Sketch of proof. Fix a symmetric counting contract in simple monitoring strategies for two players. Arrange the team of n players around a circle, and assign each player to monitor the teammate to her right, using the same p, F, s, and v as for two players. This yields a contract in simple and symmetric strategies in which the IC constraints are unchanged. Given the scaling of benefits for the larger team, the $n = 2$ and $n > 2$ optimization problems share the same objective function.

Conversely, fix a symmetric counting contract in simple and symmetric monitoring strategies with $n > 2$. Select any two players and assign them to monitor only each other, but in the same amount as in the $n > 2$ equilibrium. Since the original strategies for $n > 2$ were simple and symmetric, once again the IC constraints are unchanged and the two problems share the same objective. □

In the class of linear contracts, our conclusions about asymmetric settings also extend naturally. First, in a multiplayer setting with symmetric players, a symmetric contract is optimal (ignoring integer issues), as can easily be seen by generalizing the graphical analysis of Section 6 to multiple dimensions. Starting from symmetry, if player 1's memory improves slightly (in terms of capacity or reliability), then player 1's task burden increases slightly, while the other players' burdens decrease slightly. Just as in the two-player case, individual rationality constraints can bind when the parameters are further away from symmetry.

7.2 Players as team leaders

Suppose there are multiple teams, and we think of each "player" i in our model as a team with M_i members. Tasks are sufficiently complex that each team member can monitor or perform at most one task. As alluded in the introduction, this interpretation formalizes the idea that resource constraints can arise from limited staffing. Since the disutility of completing each task accrues to a different team member (rather than to the team), each single-task incentive constraint must be satisfied separately. Forgiving contracts are not helpful in this setting, since the individual who is forgiven for failing to complete task X is not the same individual who completed task Y. So when each "player" is a team leader, it makes sense to restrict attention to linear contracts.

If there are several types of agents with different recall abilities, we can associate each team i with recall parameter λ_i. Under this interpretation, the $\overline{\text{IR}}$ constraints studied in Section 6 should be adapted slightly: those players assigned to perform tasks should either expect to benefit enough from the tasks of others to outweigh their costs of effort, or be paid a fixed amount ex ante. In principle, this interpretation does not allow agents on the same team to monitor each other. However, there may be multiple teams with the same type, so the restriction to monitoring only across teams is without loss of generality.

7.3 Organizational structure

The extension to n-player teams raises interesting questions about optimal organizational structure. Suppose that a principal can hire a team of n agents to perform tasks and monitor each other. The principal reaps the entire benefit B from each task, but cannot observe who performed it. He makes each agent a take-it-or-leave-it offer comprising:

1. A fixed ex ante payment, no less than zero (agents have limited liability);

2. A payment $b \geq 0$ for each task completed by the team;

3. An informal contract of punishments v, which will be implemented by the team members in equilibrium (punishments destroy surplus; they are not transfers to the principal);

4. A selected equilibrium in the game induced by the contract.

Each agent accepts the offer if and only if her expected utility from the offer is at least as high as her exogenous outside option. In this setting, the agents' incentive constraints are simply Eq. 3. The principal's optimization problem is similar to Eq. 4, except that $2b$ is replaced by $B - nb$ (the principal's net profit per task), and that b is a choice variable rather than a parameter.

Were the fixed ex ante payment not bounded below by zero, the principal's optimal contract would set $b \geq c$ to make the agents willing to perform the tasks without any wasteful punishments, and then extract the surplus by asking them to pay the principal ex ante for access to the tasks. But because the principal must pay the team nb for each task, if $b \geq c$ the limited liability constraint will prevent the principal from extracting all the surplus whenever $B < nc$. The informal contract v thus serves as a costly mechanism for the principal to extract more of the surplus even when the agents have limited liability. In an optimal symmetric contract when $B < nc$, the principal pays the agents nothing ex ante (he would rather increase b than the ex ante payment), and chooses b and v to maximize expected profit, subject to the constraint that the agents must be willing to accept the offer.

The principal's optimal contract gives each agent a broadly-based bonus payment b, tied to the performance of the team, which is just large enough to meet the agents' individual rationality constraints. Given b, the contract must be optimal in the sense of Eq. 4, and hence, if the probability of recall is neither too high nor too low, must be forgiving. To implement the contract, the principal does not need a finely varying set of subtle punishments; he needs a few large punishments—or maybe just one—that will be realized only infrequently. Typically, the harshest punishment an employer can visit on an employee is firing, which is enough to implement a maximally forgiving contract. These characteristics are consistent with the stylized facts of Baker et al. (1987) (individual financial incentives are rare), Oyer and Schaefer (2005) (broad-based group incentives are common), and Lazear and Shaw (2007) (teams are common) discussed in the introduction.

However, if the principal may make different offers to different agents, he may be able to improve over the symmetric contract by assigning one or more players to be supervisors who specialize purely in monitoring. These supervisors do not need to be paid for completed tasks. Instead, they can be compensated for their opportunity costs with ex ante payments. This allows the principal to pay the team just $\hat{n}b$ per task, where $\hat{n}$ is the number of non-supervisory (promise-making) agents.

Since increasing b is less costly when some agents are supervisors, the non-supervisory agents can be induced to complete more tasks. When the agents are observably heterogeneous, the principal will select his supervisors endogenously based on their memory abilities. For the special case in which the principal is restricted to offer linear contracts, the results of Section 6 imply that supervisors should be drawn from among those players with the weakest memories. That is, the principal should follow the Dilbert Principle.

8 Conclusion

We study a team setting where forgetful players with limited memories have costly but socially efficient tasks to complete and characterize optimal contracts when the team's collective memory serves as a costly monitoring device. In accordance with stylized facts from the workforce, our results suggest that the optimal punishment scheme is forgiving when only a small number of offenses are recorded. In our model, this implies that individuals in teams make promises they do not necessarily intend to keep and that their teammates take these promises with a grain of salt. Our model provides a simple formulation for studying some of the basic tradeoffs that arise when moral hazard and capacity constraints intersect, and helps explain why real-world contracts (both formal and informal) may be forgiving.

Our conclusions about asymmetric linear contracts can be viewed as endogenously allocating supervisory responsibility to those agents with the least effective memories. Intuitively, less able players have a comparative advantage in monitoring, because their low ability is more costly in production. Furthermore, our results for general counting contracts can be applied even when a vertical supervision structure is imposed exogenously, to show that empty promises and forgiving contracts can be optimal. Our framework could be extended to study how to select between horizontal and vertical supervision structures.

We assumed that memory abilities are common knowledge, but it would be interesting to consider the case in which abilities are private information at the outset. Since the more able player is optimally given a higher workload, incentives for truthful revelation would have to be built into the contract. If ex ante transfers are feasible, a more able player should be willing to accept a harsher schedule of punishments in return for a larger ex ante payment. This intuition suggests that private information introduces a substantial friction, since the harsh punishments that able players must accept to prove their abilities will occur with positive probability.

It would also be interesting allowing the memory bound to adjust endogenously to the complexity of the information being memorized. Cowan (2000), among others, suggests that the number of effective slots in memory decreases in the complexity of the information stored. If the agents can record some of the details of their tasks and then refer to their records when performing the tasks, they may have to less to memorize for each task. That is, the memory bound can be relaxed at the cost of creating and using physical records, such as less incomplete contracts. This tradeoff can be used to characterize the optimal level of contractual detail.

A Appendix: Proofs

A.1 Proof for Section 4

Proof of Theorem 1. First we verify that it is incentive compatible for player i to fulfill all the promises she recalls. The incentive constraint for player i to complete promise $(x,i) \in \pi_i \cap m_i$ is

$$b - c \geq \lambda\, \mu_{-i}\big((x,i);\pi\big)v = \lambda \min\left\{\frac{M-p}{p}, 1\right\} v, \tag{10}$$

where $\mu_{-i}\big((x,j);\pi\big)$ denotes the marginal probability that $\mu_{-i}(\pi)$ assigns to (x,j). This constraint is guaranteed by the condition $\frac{1}{2}M \leq p \leq \frac{\lambda v}{b-c+\lambda v}M$, which in turn is implied by the conditions on λ and p in the theorem.

Next we demonstrate that either $b - c = \frac{M-p}{p}\lambda v$ or $p = 0$. If $0 < p < \frac{1}{2}M$ and the incentive constraints are satisfied, then in the promise-making stage each player can memorize all of his teammate's tasks with probability 1 and still have at least two empty slots left over, so each player can promise an additional task for which the incentive constraint is also satisfied.[19] Hence in any optimal equilibrium in which $p^* > 0$, we must have $p \geq \frac{1}{2}M$. Therefore, assuming $p > 0$, we can simplify each incentive constraint to $b - c \geq \frac{M-p}{p}\lambda v$, or, equivalently, $p \leq \frac{\lambda v}{b-c+\lambda v}M$. However, if this constraint is slack, then it would improve matters to marginally increase v, reducing the severity of punishments (which occur with positive probability) without disrupting any incentive constraints. Hence either $b - c = \frac{M-p}{p}\lambda v$ or $p = 0$.

Now we consider the problem of choosing p and v optimally. Clearly, if $p = 0$ then it is optimal to set $v = 0$, attaining zero utility for both players. So suppose that $p > 0$; then an optimal contract solves

$$\max_{p \in \mathbb{I}, v \in [\underline{v}, 0]} 2p\left(\lambda(2b-c) + (1-\lambda)\lambda\frac{M-p}{p}v\right) \qquad \text{s.t.} \quad \frac{1}{2}M \leq p \leq \frac{\lambda v}{b-c+\lambda v}M. \tag{11}$$

Since the incentive constraints bind, it suffices to solve

$$\max_{p \in \mathbb{I}} 2p\big(\lambda(2b-c) + (1-\lambda)(b-c)\big) \qquad \text{s.t.} \quad \frac{1}{2}M \leq p \leq \frac{\lambda \underline{v}}{b-c+\lambda \underline{v}}M. \tag{12}$$

[19] Here we use the assumption that M is even. If M were odd, an optimal symmetric contract might leave the one leftover slot empty, but there would be a superior asymmetric contract in which one player uses the leftover slot to make an extra promise and the other player uses it for monitoring.

Clearly $\lambda \geq \frac{b-c}{\underline{v}}$ is a necessary condition for this problem to have a solution. Since the objective and the constraints are linear in p, it is easy to see that for $\lambda \geq \max\{\frac{b-c}{\underline{v}}, \frac{c-b}{b}\}$ it is optimal to maximize p subject to the constraints; i.e., set $p = \lfloor \frac{\lambda \underline{v}}{b-c+\lambda \underline{v}} M \rfloor$ and $v = \frac{p(b-c)}{\lambda(M-p)}$. In contrast, for $\lambda < \max\{\frac{b-c}{\underline{v}}, \frac{c-b}{b}\}$ the players cannot earn positive utility from this problem (if it has a solution), so it is optimal to set $p = v = 0$. $\square$

A.2 Proofs for Section 5

Let $t_s(a) = \sum_{a'=a}^{p} \mathbb{I}\big(s(a') = a\big)\binom{p}{a'}\lambda^{a'}(1-\lambda)^{p-a'}$ denote the probability of completing a tasks given task-completion strategy s. Let $h_v(a) \equiv \sum_{f=0}^{F} v(f)g(f,a)$ be the expected punishment for fulfilling a promises given a contract v. We write $h(a)$ whenever v is implicitly known.

Lemma 1. *The value of an optimal contract in simple monitoring strategies (Eq. 4) is continuous in λ. The correspondence mapping λ to the set of optimal contracts in simple monitoring strategies, using strategies of the form $s(k)$, is upper hemicontinous.*

Proof. By Berge's Theorem of the Maximum (e.g., Aliprantis and Border 2006, Theorem 17.31). $\square$

Lemma 2 (Only deserved punishments)**.** *In any optimal contract, $v(0) = 0$.*

Proof. In an optimal contract, the upward incentive constraints in Eq. 3 can be dropped as discussed earlier. Because $g(0,a)$ is decreasing in a, the downward incentive constraints can only be relaxed by imposing $v(0) = 0$. $\square$

Proof of Theorem 2. By incentive-compatibility, to ensure that a rather than $a-1$ promises are fulfilled when a are recalled, we need $h_v(a-1) \leq h_v(a) + b - c$. By induction, $h_v(a) \leq h_v(p) + (p-a)(b-c)$, with $h_v(p) = 0$ by Lemma 2. Letting $v(f) = f\frac{p}{\lambda F}(b-c)$,

$$h_v(a) = \sum_{f=0}^{F} v(f)g(f,a) = \frac{p}{\lambda F}(b-c)\sum_{f=0}^{F} f g(f,a) = (p-a)(b-c)$$

because the expectation of the compound hypergeometric-binomial is $(p-a)\lambda\frac{F}{p}$. Moreover, this contract gives expected social utility

$$\begin{aligned}
&2\sum_{a=0}^{p}\binom{p}{a}\lambda^a(1-\lambda)^{p-a}[(2b-c)a + (p-a)(b-c)] \\
&= 2p(b-c)\sum_{a=0}^{p}\binom{p}{a}\lambda^a(1-\lambda)^{p-a} + 2b\sum_{a=0}^{p} a\binom{p}{a}\lambda^a(1-\lambda)^{p-a} \\
&= 2p(b-c+\lambda b).
\end{aligned}$$

This is positive if $\lambda > \frac{c-b}{b}$ and largest for $p = M-1$. We now prove a linear contract is optimal when λ is sufficiently high or low:

(i) At $\lambda = 1$, in every optimal contract each player must promise $p = M-1$ tasks and fulfill all of them ($s(M-1) = M-1$); the contract must impose severe enough punishments to make it incentive compatible for them to do so, but the punishments may be arbitrarily severe since they are not realized on the equilibrium path. The value of any such contract is $2(M-1)(2b-c)$. For $\lambda \to 1$, by Lemma 1 the value of the contract must converge to $2(M-1)(2b-c)$, and so must satisfy $p = M-1$ and $s(M-1) = M-1$ for λ sufficiently high. To minimize the cost of punishments, all the downward constraints $s(M-1)$ should bind, which is achieved by a linear contract. Finally, given a linear contract, $s(k) = k$ for all k is optimal.

(ii) At $\lambda = 0$, in any optimal contract either $s(k) = 0$ for all k or $v(f) = 0$ for all f. As $\lambda \to 0$, by Lemma 1 the optimal contracts must converge to either $s(k) = 0$ for all k or $v(f) = 0$ for all f. If punishments converge to zero, then it is incentive compatible only for the players to choose $s(k) = 0$ for all k, in which case it is optimal to set the punishments to exactly $v(f) = 0$ for all f. If the strategies converge to anything other than $s(k) = 0$ for all k, then for incentive compatibility the punishments must diverge ($v(f) \to -\infty$ for some f)—but the value of such contracts does not converge to zero, contrary to Lemma 1. Hence for λ sufficiently low, $s(k) = 0$ for all k and $v(f) = 0$ for all f. □

Lemma 3. *For all M, there exists $\alpha_M \in (1,2)$ such that if $c < b\alpha_M$, there exists $\tilde{\lambda} > \frac{c-b}{b}$ so that for all $\lambda \in (\frac{c-b}{b}, \tilde{\lambda})$, making and keeping $M-1$ promises yields positive social utility, but is dominated by making $\lfloor \frac{M+1}{2} \rfloor$ and employing a cutoff strategy with $p^* = 1$.*

Proof. Suppose for simplicity that M is odd and let $p = \frac{M+1}{2}$, and $F = \frac{M-1}{2}$.[20] Consider implementing the strategy where exactly one task is accomplished whenever at least one is remembered. Set $v(0) = v(1) = \cdots = v(F-1) = 0$. This implies $h(a) = 0$ for all $a > 1$.

For doing just one task to be incentive compatible, it must be that $h(1) - h(0) \geq c - b$ and $h(a) - h(1) \leq (c-b)(a-1)$ for all $a \in \{2, 3, \ldots, p\}$. For the latter condition, it suffices that $h(1) \geq b - c$. For the latter condition, observe that $h(1) = v(F)g(F,1)$ and $h(0) = h(1)\frac{g(F,0)}{g(F,1)}$. Since

$$\frac{g(F,0)}{g(F,1)} = \frac{\binom{p}{F}}{\binom{p-1}{F}} = \frac{p}{p-F},$$

$h(0) = \frac{p}{p-F}h(1)$. Therefore, IC requires $h(1) \leq \frac{p-F}{F}(b-c)$. Let us set $h(1) = \frac{2}{M-1}(b-c)$ and $h(0) = \frac{M+1}{M-1}(b-c)$.

[20] The even memory proof is essentially identical: use the scheme here, under-utilizing one slot, and compare to the full-utilization linear contract.

Therefore this contract is feasible and incentive compatible, and has expected social utility

$$2\Big[\big(1-(1-\lambda)^{\frac{M+1}{2}}\big)\Big(\frac{2(b-c)}{M-1}+2b-c\Big)+(1-\lambda)^{\frac{M+1}{2}}(b-c)\frac{M+1}{M-1}\Big].$$

After some algebra, this expression is larger than $2(M-1)(b-c+b\lambda)$ (the expected social utility from the optimal contract implementing $M-1$ promises and fulfilling all those remembered) if

$$\frac{c(M^2-3M)-b(M^2-4M+1)}{b(M-1)} > (1-\lambda)^{\frac{M+1}{2}}+(M-1)\lambda. \tag{13}$$

Define $\phi:[0,1]\to\mathbb{R}$ by $\phi(\lambda)=(1-\lambda)^{\frac{M+1}{2}}+(M-1)\lambda$, and note that ϕ is strictly increasing. Let

$$\bar{\lambda}=\phi^{-1}\Big(\frac{c(M^2-3M)-b(M^2-4M+1)}{b(M-1)}\Big).$$

To show that Eq. 13 holds for $\lambda\in(\frac{c-b}{b},\bar{\lambda})$, it suffices to show that $\frac{c-b}{b}<\bar{\lambda}$, or that

$$\frac{c(M^2-3M)-b(M^2-4M+1)}{b(M-1)} > \phi(\frac{c-b}{b}).$$

After some algebra, this holds if

$$(2-\frac{c}{b})^{\frac{M+1}{2}} < \frac{2M}{M-1}-\frac{c}{b}\frac{M+1}{M-1}.$$

Define $\hat{\phi}:[1,2]\to\mathbb{R}$ by

$$\hat{\phi}(x)=\frac{2M}{M-1}-x\frac{M+1}{M-1}-(2-x)^{\frac{M+1}{2}}.$$

It can be seen that $\hat{\phi}$ is concave, first increasing and eventually negative, with a unique $\alpha(M)\in(1,2)$ such that $\hat{\phi}(\alpha(M))=0$. Hence the bound $c<b\alpha(M)$. □

For the following lemma, we say that a function $\psi:\{0,1,\dots,R\}\to\mathbb{R}$ is *concave* if $\psi(r+1)-\psi(r)\le\psi(r)-\psi(r-1)$ for all $r=1,\dots,R-1$. A function $\phi:\mathcal{Z}\to\mathbb{R}$, where $\mathcal{Z}\subseteq\mathbb{R}$, is *double crossing* if there is a (possibly empty) convex set $A\subset\mathbb{R}$ such that $A\cap\mathcal{Z}=\big\{z\in\mathcal{Z}:\phi(z)<0\big\}$.

Lemma 4. *Let $\mathcal{R}=\{0,1,\dots,R\}$, and let $\{q_z\}_{z\in\mathcal{Z}}$ be a collection of probability distributions on $\mathcal{R}$ parameterized by z, which takes either discrete values $z\in\mathcal{Z}=\{0,1,\dots,Z\}$ or continuous values $z\in\mathcal{Z}=[0,1]$. If*

1. *There exists $k,c\in\mathbb{R}$, $k\neq 0$, such that $z=k\sum_{r=0}^{R} rq_z(r)+c$ for all $z\in\mathcal{Z}$;*
2. *Either $q_{z+1}(r)-2q_z(r)+q_{z-1}(r)$ (for all $z=1,\dots,Z-1$ if z is discrete) or $\frac{\partial^2}{\partial z^2}q_z(r)$ (for all $z\in(0,1)$ if z is continuous), as a function of r, is double crossing;*

3. $\psi : \{0, 1, \ldots, R\} \to \mathbb{R}$ *is concave;*

then $\Psi(z) = \sum_{r=0}^{R} \psi(r) q_z(r)$ *is concave.*[21]

Proof. Since $z = k \sum_{r=0}^{R} r q_z(r) + c$, there exists $\hat{b} \in \mathbb{R}$ such that $\sum_{r=0}^{R}(mr + b) q_z(r) = \frac{m}{k} z + \hat{b} + c$ for any real m and b. Hence, for any m and b, if z is discrete then

$$\sum_{r=0}^{R}(mr + b)\big(q_{z+1}(r) - 2q_z(r) + q_{z-1}(r)\big) = \frac{m}{k}\big(z + 1 - 2z + z - 1\big) = 0,$$

for all $z = 1, \ldots, Z - 1$, while if z is continuous then for all $z \in (0, 1)$,

$$\sum_{r=0}^{R}(mr + b)\frac{\partial^2}{\partial z^2} q_z(r) = \frac{\partial^2}{\partial z^2}\Big(\frac{m}{k} z + \hat{b} + c\Big) = 0.$$

Therefore, for any m and b, if z is discrete, the second difference of $\Psi(z)$ can be written as

$$\begin{aligned} \Psi(z + 1) - 2\Psi(z) + \Psi(z - 1) &= \sum_{r=0}^{R} \psi(r)\big(q_{z+1}(r) - 2q_z(r) + q_{z-1}(r)\big) \\ &= \sum_{r=0}^{R}\big(\psi(r) - mr - b\big)\big(q_{z+1}(r) - 2q_z(r) + q_{z-1}(r)\big), \end{aligned} \tag{14}$$

while if z is continuous then the second derivative of $\Psi(z)$ can be written as

$$\frac{\partial^2}{\partial z^2}\Psi(z) = \sum_{r=0}^{R} \psi(r)\frac{\partial^2}{\partial z^2} q_z(r)\big) = \sum_{r=0}^{R}\big(\psi(r) - mr - b\big)\frac{\partial^2}{\partial z^2} q_z(r). \tag{15}$$

By assumption, either $q_{z+1}(r) - 2q_z(r) + q_{z-1}(r)$ (if z is discrete) or $\frac{\partial^2}{\partial z^2} q_z(r)$ (if z is continuous), as a function of r, is double crossing. Furthermore, since ψ is concave, we can choose m and b such that, wherever $\big(q_{z+1}(r) - 2q_z(r) + q_{z-1}(r)\big)$ or $\frac{\partial^2}{\partial z^2} q_z(r)$ is nonzero, $\psi(r) - mr - b$ either has the opposite sign or is zero. From Eq. 14 and Eq. 15, above, we may conclude $\Psi(z)$ is concave. □

Lemma 5. *For some strategy* s*, suppose that* p^* *satisfies* $p - (p^* - 1) \geq F$ *and that*

$$\sum_{a=0}^{p} t_s(a)\left(g(f, a) - g(F, a)\frac{g(f, p^*) - g(f, p^* - 1)}{g(F, p^*) - g(F, p^* - 1)}\right) \geq 0 \textit{ for all } f = 1, \ldots, F - 1. \tag{16}$$

Then the contract is suboptimal if it does not involve cutoff strategies. Moreover, the best-case

[21] A more general mathematical result along these lines appears in Fishburn (1982), but the condition there is not as easy to check.

punishments for implementing a cutoff strategy p^* are given by

$$\frac{c-b}{g(F,p^*)-g(F,p^*-1)}\Big(\sum_{a=0}^{p^*-1}\binom{p}{a}\lambda^a(1-\lambda)^{p-a}g(F,a)+g(F,p^*)\sum_{a=p^*}^{p}\binom{p}{a}\lambda^a(1-\lambda)^{p-a}\Big), \qquad (17)$$

derived by setting $v(f)=0$ for $f<F$ and $v(F)$ high enough to make p^* indifferent to p^*-1.

Proof. If a contract is optimal, we can ignore the upward incentive constraints (if any bind, then it would be optimal to do that number of tasks). Suppose that s is optimal given p, F and is not a cutoff strategy. Fixing s, finding the optimal punishments is a linear programming problem. By duality theory, we know that if the primal problem is $\max u^T y$ s.t. $A^T y \leq w$ and $y \geq 0$, then the dual problem is $\min w^T x$ s.t. $Ax \geq u$ and $x \geq 0$; the optimal solution to one problem corresponds to the Lagrange multipliers of the other, and if feasible solutions to the dual and primal achieve the same objective value then these are optimal for their respective problems.

The relaxed problem (dropping upward incentive constraints), written in the form of the primal problem, is given by

$$\begin{aligned}
&\max \sum_{f=0}^{F}(-v(f))\sum_{a=0}^{p}-g(f,a)t_s(a) \text{ subject to}\\
&\sum_{f=0}^{F}(-v(f))[g(f,a)-g(f,k)] \leq -(a-k)(c-b) \text{ for all } a \text{ s.t. } t_s(a)>0 \text{ and all } k<a\\
&\text{and } -v(f)\geq 0 \text{ for all } f=0,1,\dots,F
\end{aligned}$$

The dual of this problem is then

$$\begin{aligned}
&\min \sum_{\{(k,a)\ \mid\ t_s(a)>0,k<a\}} -(a-k)(c-b)x_{ka} \text{ subject to}\\
&\sum_{\{(k,a)\ \mid\ t_s(a)>0,k<a\}} x_{ka}[g(f,a)-g(f,k)] \geq -\sum_{a=0}^{p} g(f,a)t_s(a) \text{ for all } f=0,1,\dots,F\\
&\text{and } x_{ka}\geq 0 \text{ for all } (k,a) \text{ s.t. } k<a,\ t_s(a)>0.
\end{aligned}$$

Let $v(f)=0$ for all $f=0,1,\dots,F-1$, and set $v(F)=\frac{c-b}{g(F,p^*)-g(F,p^*-1)}$, which makes the IC constraint bind in comparing p^* and p^*-1 tasks. We know the denominator is strictly negative by the assumption that $p-(p^*-1)\geq F$ and the fact that $g(F,a)\leq g(F,a-1)$ for all $a=1,2,\dots,p$. This is feasible in the primal because all downward IC constraints will be slack after the first that binds, by preservation of convexity in Lemma 4. Then the value of this solution to the primal is given by

$$\frac{c-b}{g(F,p^*)-g(F,p^*-1)}\sum_{a=0}^{p} g(F,a)t_s(a).$$

Let $x_{ka}=0$ for all pairs (k,a) except for $a=p^*$ and $k=p^*-1$, since those IC constraints in the primal are slack. Let

$$x_{p^*,p^*-1}=-\frac{\sum_{a=0}^{p} g(F,a)t_s(a)}{g(F,p^*)-g(F,p^*-1)},$$

corresponding to the constraint for F binding, since $v(F)<0$. This is feasible in the dual by the assumption in (16). Then the value of the dual is the same as that in the primal, which means that the optimal punishment involves $v(f)=0$ for all $f=0,1,\ldots,F-1$ and $v(F)=\frac{c-b}{g(F,p^*)-g(F,p^*-1)}$.

However, because all downward IC constraints are satisfied, if s is not a cutoff strategy then at least one of the upward IC constraints that were dropped is violated, a contradiction to being an optimal strategy given p and F. □

Proof of Theorem 3. The optimality of empty promises and part (i) are proved in Lemma 3, which shows promise-keeping with $p \leq M-1$ is dominated in this range. To see part (ii), note that by Theorem 2, a linear contract can be optimal only if $p=M-1$ (because punishments are not bounded below). We now show that the converse holds. Fix $p=M-1$; then $F=1$ and the hypotheses of Lemma 5 are satisfied, so the optimal contract uses cutoff strategies. Suppose to the contrary of the converse that $0<p^*<p=M-1$ is optimal. Then by Lemma 5, $v(0)=0$ and $v(1)$ is set to make doing p^* tasks indifferent to doing p^*-1 tasks: that is, $v(F)=\frac{c-b}{g(F,p^*)-g(F,p^*-1)}$. Then the expected punishment when a tasks are done is given by

$$(c-b)\frac{g(F,a)}{g(F,p^*)-g(F,p^*-1)}=(c-b)\frac{\binom{M-1-a}{1}}{\binom{M-1-p^*}{1}-\binom{M-p^*}{1}}=-(c-b)(M-1-a).$$

Consequently, expected punishment is independent of p^*, and decreases in a. Because benefits are also increasing in a, the contract is dominated by cutoff strategies with $p^*=M-1$; i.e., a linear contract. Promise-keeping, in turn, is dominated by not making any promises if $\lambda<\frac{c-b}{b}$. □

Lemma 6. *For memory capacity M, promises p, cutoff p^*, and monitoring slots F satisfying $0<p^*\leq p<M$ and $0<F\leq p-p^*+1$, the optimal social welfare from implementing a p^*-cutoff strategy s using a maximally forgiving contract is*

$$2\sum_{a=0}^{p}\binom{p}{a}\lambda^a(1-\lambda)^{p-a}\Big((2b-c)s(a)+\frac{(c-b)g\big(F,s(a)\big)}{g(F,p^*)-g(F,p^*-1)}\Big). \tag{18}$$

Moreover,

1. *The value of Eq. 18 is strictly increasing and concave in λ.*

2. *If $p < M-1$ and $p_1^* < p_2^* \le p-F+1$, the value of Eq. 18 for p_2^* strictly single crosses the value of Eq. 18 for p_1^* from below, as functions of λ.*

3. *If $z \in \mathbf{Z}_{++}$, $p+z \le M-F$, and $p^* \le p-F+1$, the value of Eq. 18 for $p+z, p^*+z$ strictly single crosses the value of Eq. 18 for p, p^* from below, as functions of λ.*

Proof. We prove each part separately below.

(i) The value of Eq. 18 is the expectation of $\beta(a) \equiv 2(2b-c)s(a) + 2\frac{(c-b)g(F,s(a))}{g(F,p^*)-g(F,p^*-1)}$ with respect to the binomial distribution over a. For any cutoff strategy s, $(2b-c)s(a)$ is clearly concave. The second term of $\beta(a)$ is a negative constant times $g(F, s(a))$. Itself, $g(F,s(a)) = \lambda^F \binom{p-s(a)}{F} \Big/ \binom{p}{F}$, which is convex:

$$\binom{p-s(a+1)}{F} - 2\binom{p-s(a)}{F} + \binom{p-s(a-1)}{F}$$
$$= \begin{cases} \binom{p-a}{F}\left(\frac{F}{p-(a+1)-F} - \frac{F}{p-a}\right) & \text{if } a \le p^*-1, \\ \binom{p-(p^*-1)}{F} - \binom{p-p^*}{F} & \text{if } a = p^*, \\ 0 & \text{if } a \ge p^*+1. \end{cases}$$

which is positive because $F \ge 1$, and $p - p^* + 1 \ge F$. Hence $\beta(a)$ is concave. Finally, the binomial distribution satisfies double-crossing, since

$$\frac{\partial^2}{\partial \lambda^2}\left(\binom{p}{a}\lambda^a(1-\lambda)^{p-a}\right) = \binom{p}{a}(1-\lambda)^{p-2-a}\lambda^{a-2}\left(a^2 - \big(1+2(p-1)\lambda\big)a + p(p-1)\lambda^2\right)$$

is negative if and only if $a^2 - \big(1+2(p-1)\lambda\big)a + p(p-1)\lambda^2 < 0$. Hence by Lemma 4, Eq. 18 is concave in λ. To see that Eq. 18 is increasing in λ, observe that the benefit of each task is linear in a, increasing in p^* and independent of λ, which is a parameter of first-order stochastic dominance for the binomial distribution.

(ii) For a cutoff strategy s, the expected punishment for completing $s(a)$ tasks is

$$\frac{(c-b)g(F,s(a))}{g(F,p^*)-g(F,p^*-1)}.$$

Since λ cancels out of the above, we need only check that has increasing differences in a and p^* (by Corollary 10 of Van Zandt and Vives 2007). Let us denote a p^*-cutoff strategy by s_{p^*}.

Since $c - b > 0$, the sign of the second difference depends on

$$\frac{g(F, s_{p^*+1}(a+1)) - g(F, s_{p^*+1}(a))}{g(F, p^*+1) - g(F, p^*)} - \frac{g(F, s_{p^*}(a+1)) - g(F, s_{p^*}(a))}{g(F, p^*) - g(F, p^*-1)}$$
$$= \begin{cases} 0 & \text{if } a \geq p^*+1 \\ 1 & \text{if } a = p^* \\ \frac{g(F,a+1)-g(F,a)}{g(F,p^*+1)-g(F,p^*)} - \frac{g(F,a+1)-g(F,a)}{g(F,p^*)-g(F,p^*-1)} & \text{if } a \leq p^*-1. \end{cases} \tag{19}$$

Concentrating on the third case, since $g(F, a)$ is decreasing in a, it suffices to show that

$$\binom{p-p^*}{F} - \binom{p-p^*+1}{F} > \binom{p-p^*+1}{F} - \binom{p-p^*+2}{F}. \tag{20}$$

But this is exactly analogous to the calculation in part (1).

(iii) Without loss, the limit of summation in Eq. 18 may be replaced with M, since $\binom{p}{a} = 0$ for $a > p$. Let $p + z'$ be the number of promises for determining the distribution of a, but let $p + z$ be the number of promises for determining $g(F, \cdot)$. For feasibility, we must have $z' = z$, but for now relax feasibility.

First we show that z is optimally increasing in (λ, z'). Similarly to above, we check that

$$\frac{\binom{p+z-s_{p^*+z}(a)}{F}}{\binom{p+z-(p^*+z)}{F} - \binom{p+z-(p^*+z-1)}{F}}$$

has increasing differences in a and z. The sign of the second difference is determined by

$$\frac{\binom{p+z+1-s_{p^*+z+1}(a+1)}{F} - \binom{p+z+1-s_{p^*+z+1}(a)}{F}}{\binom{p-p^*}{F} - \binom{p-p^*+1}{F}} - \frac{\binom{p+z-s_{p^*+z}(a+1)}{F} - \binom{p+z-s_{p^*+z}(a)}{F}}{\binom{p-p^*}{F} - \binom{p-p^*+1}{F}}$$
$$= \begin{cases} 0 & \text{if } a \geq p^*+z+1 \\ 1 & \text{if } a = p^*+z \\ \frac{\binom{p+z-a}{F} - \binom{p+z-a+1}{F} - \binom{p+z-1-a}{F} + \binom{p+z-a}{F}}{\binom{p-p^*}{F} - \binom{p-p^*+1}{F}} & \text{if } a \leq p^*+z-1. \end{cases}$$

The case $a \leq p^*+z-1$ reduces to checking the numerator is negative, since the denominator is negative. Again, by substituting in the definition of the binomial coefficient, this is equivalent to $\frac{p+z-a+1-F}{p+z-a} \leq 1$, which holds because $F \geq 1$ and $p + z - a + 1 > 0$.

Then the result derives from observing that for feasible contracts z', which is also a parameter of stochastic dominance for the binomial distribution, is itself increasing in z. Hence the optimal choice of z is increasing in λ (i.e., single crossing as claimed). $\square$

Proof of Theorem 4. Recall that task-completion strategies are necessarily increasing step functions which jump to the maximum whenever a jump occurs. The first nontrivial case is $M = 3$, in which the only possible optimal promise levels are then $p = 1$ (with $F = 1$) and $p = 2$ (with $F = 1$). In both cases Theorem 3 implies the contract must be promise-keeping.

For the case $M = 4$, the only possible optimal promise levels are $p = 2$ (with $F = 2$) and $p = 3$ (with $F = 1$). Theorem 3 implies that the last case again reduces to promise-keeping with linear contracts, and that $p = 2$ (with $F = 2$) is suboptimal unless it is a cutoff strategy with $p^* = 1$. In this case the assumptions of Lemma 5 are satisfied, because for the case $F = 2$ the part of the summand in Eq. 16 that is in parentheses is always nonnegative, for all $\lambda \in (0, 1)$ and choices of p, p^*, a that are feasible given that $M \leq 5$.[22]

Finally, for the case $M = 5$, the only possible optimal promise levels are $p = 2$ (with $F = 2$), $p = 3$ (with $F = 2$), and $p = 4$ (with $F = 1$). The last case again reduces to promise-keeping with linear contracts by Theorem 3. Strategies must be weakly increasing for the contract to be optimal, and by Theorem 3, they cannot have empty promises if $p^* = p$. Then there is only a cutoff strategy remaining for $p = 2$, with $p^* = 1$ (same as for $M = 4$). Moreover, there is only one non-cutoff strategy for the case that $p = 3$ that could potentially be optimal: $s(a) = 0$ for $a < 2$, and $s(a) = 2$ for $a \geq 2$. To rule this out, observe that the assumptions in Lemma 5 are satisfied for $p = 3$ and $M = 5$, so a non-cutoff strategy cannot be optimal. The cutoff strategies (p, p^*) remaining are given by $(x, 0)$, $(2, 1)$, $(3, 1)$, $(3, 2)$, and $(4, 4)$ are potentially optimal. We know by the single crossing result for fixed $p = 3$ that $(3, 2)$ single crosses $(3, 1)$ from below, and also single crosses $(2, 1)$ from below. By Lemma 6 the value functions for each p^* are concave in λ, so that once the linear value function for $p = 4$ is optimal it remains so. $\square$

Lemma 7. *If v is decreasing convex, then $h \equiv \sum_{f=0}^{F} v(f)g(f, \cdot)$ is decreasing convex.*

Proof. Note that decreasing convex is literally concave. By reversing the order of summation, and the fact that $\binom{k}{f} = 0$ when $k < f$, we can write $h(a)$ as follows:

$$\begin{aligned}
\sum_{f=0}^{F} g(f, a)v(f) = & \sum_{f=0}^{F} \left(\sum_{k=0}^{F} \frac{\binom{p-a}{k}\binom{a}{F-k}}{\binom{p}{F}} \binom{k}{f} \lambda^f (1-\lambda)^{k-f} \right) v(f) \\
= & \sum_{k=0}^{F} \frac{\binom{p-a}{k}\binom{a}{F-k}}{\binom{p}{F}} \left(\sum_{f=0}^{F} \binom{k}{f} \lambda^f (1-\lambda)^{k-f} v(f) \right).
\end{aligned}$$

Therefore, the expectation is first with respect to the binomial, and then with respect to the hypergeometric. Using Lemma 4 twice gives the result. First, note that the expectation of the binomial is λk, a linear function of k, while the expectation of the hypergeometric is $\frac{F}{p}(p - a)$, a linear function of a. Hence it suffices to show that the binomial second-difference in k is double-

[22] This is tedious to check analytically but easily checked numerically.

crossing in f (hence the inside expectation is decreasing convex in k) and the hypergeometric second-difference in a is double-crossing in k. To see this is true for the binomial, note that we may write the binomial second-difference in k as

$$\binom{k}{f}\lambda^f(1-\lambda)^{k-f}\left(\frac{(k+1)(1-\lambda)}{k+1-f} - 2 + \frac{k-f}{k(1-\lambda)}\right).$$

It can be shown that the term in parentheses is strictly convex in f and therefore double crossing in f, so the whole expression is double-crossing in f. To see this is true for the hypergeometric, note that we may write the hypergeometric second-difference in a as

$$\frac{\binom{p-a}{k}\binom{a}{F-k}}{\binom{p}{F}}\left(\frac{p-a-k}{p-a}\cdot\frac{a+1}{a+1-F+k} - 2 + \frac{p-a+1}{p-a+1-k}\cdot\frac{a-F+k}{a}\right).$$

It can be shown that the term in parentheses has either no real roots or exactly two real roots.[23] If there are no real roots, then the term in parentheses is double-crossing in k (recall that the region in which it is negative must be convex, but may be empty), and therefore the whole expression is double-crossing in k. If there are two real roots, it can be shown that the derivative with respect to k is negative at the smaller root, and that therefore both the term in parentheses and the whole expression are double-crossing in k. □

Proof of Theorem 5. Fix any p, F, λ. Suppose that the strategy s, with $p^* > 0$ the maximal number of tasks completed, is optimal. Consider the decreasing convex contract v that implements s at minimum cost. Because v is decreasing, MLRP (more weakly, FOSD in a) implies expected punishments are decreasing in the number of tasks completed: $h(a) > h(a-1)$ for all a. By contradiction, suppose that the downward constraint for p^* versus p^*-1 is slack: $h(p^*)-h(p^*-1) > c-b$. By Lemma 7 and monotonicity, for any $k > 1$, $h(p^*-k+1) - h(p^*-k) > c-b$. But then for every a such that $s(a) = a$, and every $a' < a$, the downward constraint $h(a) - h(a') = \sum_{k=a'}^{a-1} h(k+1) - h(k) \geq (a-a')(c-b)$ must be slack. However, some constraint must bind at the optimum, else the strategy is implementable for free, so it must be that the downward constraint for p^* versus p^*-1 binds. Again, each downward constraint is satisfied, and for any $a > p^*$, $h(a) - h(p^*) < (a-p^*)(c-b)$. So the strategy s is a p^*-cutoff.

Now, suppose that we look for the optimal convex contract implementing p promises , F monitoring slots, and cutoff strategy s with cutoff p^*. By the argument above, the only incentive constraint that binds is the downward constraint for completing p^* promises. Since $v(0) = 0$ (by Lemma 2), convexity implies monotonicity. Furthermore the constraint $v(0) \geq 0$ does not bind,[24]

[23] The term in parentheses does not account for the fact that the entire expression equals zero whenever $k > p-a$ or $F-k > a$. However, on the closure of these regions the second difference cannot be negative, and so these regions may be ignored.

[24] Although it is satisfied with equality, by Lemma 2 relaxing the constraint would not change the solution.

so the cost minimization problem (dropping all other incentive constraints) in primal form is

$$\max_{(-v)\geq\vec{0}} \sum_{f=0}^{F}\Big(-(-v(f))\sum_{a=0}^{p} -g(f,a)t_s(a)\Big) \text{ subject to}$$
$$\sum_{f=0}^{F}(-v(f))[g(f,p^*)-g(f,p^*-1)] \leq -(c-b),$$
$$2(-v(f))-(-v(f+1))-(-v(f-1)) \leq 0 \text{ for all } f=1,\ldots,F-1.$$

Let x be the Lagrange multiplier for the lone incentive compatibility constraint, z_f the Lagrange multiplier for the convexity constraint $2(-v(f))-(-v(f+1))-(-v(f-1)) \leq 0$, and $\vec{z}$ the vector $(z_1,\ldots,z_{F-1})$. The constraint set for this problem can then be written as $A^\top \cdot (-v(0),\ldots,-v(F))$, where, in sparse form,

$$A = \begin{pmatrix} g(0,p^*)-g(0,p^*-1) & -1 & & & \\ \vdots & 2 & \ddots & & \\ \vdots & -1 & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & \ddots & \\ \vdots & & & \ddots & \ddots & -1 \\ \vdots & & & & \ddots & 2 \\ g(F,p^*)-g(F,p^*-1) & & & & & -1 \end{pmatrix}.$$

Let r be the vector of dual variables: $r=(x,z_1,\ldots,z_{F-1})$. The dual problem is

$$\min_{r\geq\vec{0}} (b-c)x \quad \text{s.t. } (Ar)_f \geq -\sum_{a=0}^{p} g(f,a)t_s(a) \text{ for all } f=0,1,\ldots,F,$$

where $(Ar)_f$ is the (f)th component of $A\cdot r$; i.e.,

$$(Ar)_f = x[g(f,p^*)-g(f,p^*-1)] - z_{f-1} + 2z_f - z_{f+1},$$

where we define $z_0 \equiv 0$, $z_F \equiv 0$, and $z_{F+1} \equiv 0$.

Let $\hat{f}$ be the smallest f such that $v(f)<0$. Then it must be that $v(f)<0$ for all $f \geq \hat{f}$, so by duality theory the constraint $(A\cdot r)_f \geq -\sum_{a=0}^{p} g(f,a)t_s(a)$ binds for all $f\geq\hat{f}$. Hence

$$x = \frac{\sum_{a=0}^{p} g(f,a)t_s(a) - z_{f-1} + 2z_f - z_{f+1}}{g(f,p^*-1)-g(f,p^*)} \text{ for all } f=\hat{f},\ldots,F. \tag{21}$$

In particular, this means that if $z_{F-1}=0$ (which is implied when $\hat{f}=F$) the optimal contract (which would have expected punishment $-x(c-b)$) has the same value as that derived in Lemma 5), completing the claim. In the remainder we assume $z_{F-1} > 0$.

Observe that the sum of the z-terms over $(A \cdot r)_{F-1}$ and $(A \cdot r)_F$ is $-z_{F-1} + (2z_{F-1} - z_{F-2}) = z_{F-1} - z_{F-2}$. Note also the corresponding sum of z-terms over $F-2$, $F-1$, and F: $-z_{F-1} + (2z_{F-1} - z_{F-2}) + (-z_{F-3} + 2z_{F-2} - z_{F-1}) = z_{F-2} - z_{F-3}$. Continuing in this manner, the sum of the z-terms in $(A \cdot r)_f$ from any $\tilde{f} \geq \hat{f}$ to F is $z_{\tilde{f}} - z_{\tilde{f}-1}$. Therefore, summing the equalities in Eq. 21 yields the following recursive system for $z_{\tilde{f}}$, $\tilde{f} = \hat{f}, \ldots, F$

$$z_{\tilde{f}} = z_{\tilde{f}-1} - \sum_{f=\tilde{f}}^{F} \sum_{a=0}^{p} g(f,a) t_s(a) + x \sum_{f=\tilde{f}}^{F} \big(g(f, p^*-1) - g(f,p^*)\big).$$

Also, by definition of $\hat{f}$ the convexity constraint is slack at $\hat{f}-1$, so $z_{\hat{f}-1} = 0$. Then induction yields, for $f' = \hat{f}, \ldots, F$,

$$z_{f'} = -\sum_{\tilde{f}=\hat{f}}^{f'} \sum_{f=\tilde{f}}^{F} \sum_{a=0}^{p} g(f,a) t_s(a) + x \sum_{\tilde{f}=\hat{f}}^{f'} \sum_{f=\tilde{f}}^{F} \big(g(f, p^*-1) - g(f,p^*)\big).$$

Plugging this equation for $f' = F$ into the binding constraint $(Ar)_F \geq -\sum_{a=0}^{p} g(F,a) t_s(a)$ provides solution for x in terms of $\hat{f}$:

$$x = \frac{\sum_{\tilde{f}=\hat{f}}^{F} \sum_{f=\tilde{f}}^{F} \sum_{a=0}^{p} g(f,a) t_s(a)}{\sum_{\tilde{f}=\hat{f}}^{F} \sum_{f=\tilde{f}}^{F} \big(g(f,p^*-1) - g(f,p^*)\big)}. \tag{22}$$

Note that for a random variable X on $\{0, \ldots, n\}$, the expectation of X is $\sum_{j=1}^{n} j \Pr(X=j)$ but this is also equal to $\sum_{j=1}^{n} \Pr(X \geq j)$. Since $\sum_{f=\tilde{f}}^{F} \sum_{a=0}^{p} g(f,a) t_s(a) = \Pr(f \geq \tilde{f})$, the numerator of Eq. 22 can be rewritten as

$$\begin{aligned}
\sum_{\tilde{f}=\hat{f}}^{F} \sum_{f=\tilde{f}}^{F} \sum_{a=0}^{p} g(f,a) t_s(a) &= \sum_{\tilde{f}=\hat{f}}^{F} \Pr(f \geq \tilde{f}) = \sum_{\tilde{f}=\hat{f}}^{F} (\tilde{f} - \hat{f} + 1) \Pr(f = \tilde{f}) \\
&= \sum_{\tilde{f}=1}^{F} (\tilde{f} - \hat{f} + 1)_+ \Pr(f = \tilde{f}) = \mathbb{E}\big[(f - \hat{f} + 1)_+\big] \equiv \mathbb{E}[\phi(\hat{f})],
\end{aligned}$$

where $(y)_+ \equiv \max\{y, 0\}$ and ϕ is the random function $\phi(\hat{f}) \equiv (f - \hat{f} + 1)_+$. In words, $\phi(\hat{f})$ is the number of discovered unfulfilled promises that exceed the threshold for punishment ($\hat{f}$). The denominator of Eq. 22 can be rewritten similarly, yielding

$$x = \frac{\mathbb{E}[\phi(\hat{f})]}{\mathbb{E}[\,\phi(\hat{f}) \mid a = p^* - 1\,] - \mathbb{E}[\,\phi(\hat{f}) \mid a = p^*\,]}. \tag{23}$$

The minimized expected punishment is $\mathbb{E}[v(f)] = (b-c)x$, and hence is implemented by the kinked-linear punishment schedule

$$v(f) = -\frac{(c-b)(f - \hat{f} + 1)_+}{\mathbb{E}[\,\phi(\hat{f}) \mid a = p^* - 1\,] - \mathbb{E}[\,\phi(\hat{f}) \mid a = p^*\,]} \text{ for all } f = 0, 1, \ldots, F. \qquad \square$$

The comparative statics of maximally forgiving punishments follow from Lemma 6

A.3 Proofs for Section 6

The proofs of Theorem 6 and Theorem 7 follow directly from the complete characterization below.

Lemma 8. *Suppose that $\lambda_i \geq \max\{\frac{b-c}{\underline{v}}, \frac{c-b}{b}\}$ for $i = 1, 2$. Then the optimal linear contract is characterized by four binding constraints: the original IC_1 and IC_2, and two additional binding constraints determined by $\frac{M_1}{M_2}$, σ_1, σ_2, and γ according to*

<table>
<tr><td></td><td></td><td colspan="3">M_1/M_2</td></tr>
<tr><td></td><td></td><td>$(0, \frac{\sigma_1\sigma_2+\gamma}{\sigma_1(1+\gamma)})$</td><td>$(\frac{\sigma_1\sigma_2+\gamma}{\sigma_1(1+\gamma)}, \frac{\sigma_2(1+\gamma)}{\sigma_1\sigma_2+\gamma})$</td><td>$(\frac{\sigma_2(1+\gamma)}{\sigma_1\sigma_2+\gamma}, \infty)$</td></tr>
<tr><td rowspan="3">$\frac{\sigma_2}{\sigma_1}\frac{\sigma_1-\gamma}{\sigma_2-\gamma}$</td><td>$(\frac{1}{\sigma_1}, \infty)$</td><td rowspan="2">$\overline{IR}_1$ and $\overline{IC}_1$</td><td colspan="2">$\overline{IR}_1$ and $\overline{IC}_2$</td></tr>
<tr><td>$(\sigma_2, \frac{1}{\sigma_1})$</td><td>$\overline{IC}_1$ and $\overline{IC}_2$</td><td rowspan="2">$\overline{IR}_2$ and $\overline{IC}_2$</td></tr>
<tr><td>$(0, \sigma_2)$</td><td colspan="2">$\overline{IR}_2$ and $\overline{IC}_1$</td></tr>
</table>

For each case, the number of promises is given by

$$\begin{aligned}
\overline{IC}_1 \text{ and } \overline{IC}_2: &\qquad p_1 = \frac{M_1 - \sigma_1 M_2}{1 - \sigma_1\sigma_2} \text{ and } p_2 = \frac{M_2 - \sigma_2 M_1}{1 - \sigma_1\sigma_2}, \\
\overline{IR}_1 \text{ and } \overline{IC}_1: &\qquad p_1 = \frac{M_2}{1+\gamma}\frac{\gamma}{\sigma_2} \text{ and } p_2 = \frac{M_2}{1+\gamma}, \\
\overline{IR}_1 \text{ and } \overline{IC}_2: &\qquad p_1 = \frac{\gamma M_1}{\sigma_1\sigma_2 + \gamma} \text{ and } p_2 = \frac{\sigma_2 M_1}{\sigma_1\sigma_2 + \gamma}, \\
\overline{IR}_2 \text{ and } \overline{IC}_1: &\qquad p_1 = \frac{\sigma_1 M_2}{\sigma_1\sigma_2 + \gamma} \text{ and } p_2 = \frac{\gamma M_2}{\sigma_1\sigma_2 + \gamma}, \\
\overline{IR}_2 \text{ and } \overline{IC}_2: &\qquad p_1 = \frac{M_1}{1+\gamma} \text{ and } p_2 = \frac{M_1\gamma\sigma_1}{1+\gamma}.
\end{aligned}$$

If $\lambda_i < \max\{\frac{b-c}{\underline{v}}, \frac{c-b}{b}\}$ for some i then the optimal contract has $p_1 = p_2 = 0$ and $v_1 = v_2 = 0$.

Proof. Define $\sigma_1 \equiv \frac{b-c}{\lambda_1 \underline{v}}$, $\sigma_2 \equiv \frac{b-c}{\lambda_2 \underline{v}}$, and $\gamma \equiv -\frac{b}{\underline{v}}$. Using this notation,

$$\begin{aligned} \overline{\mathrm{IC}}_1 \;&\Leftrightarrow\; p_2 \le M_2 - p_1\sigma_2 \text{ whenever } p_1 \ge M_2 - p_2, \\ \overline{\mathrm{IC}}_2 \;&\Leftrightarrow\; p_2 \le \frac{1}{\sigma_1}(M_1 - p_1) \text{ whenever } p_2 \ge M_1 - p_1. \end{aligned}$$

Under the assumption that $\lambda_i \ge \frac{b-c}{\underline{v}}$ we know $\sigma_i \in (0,1)$ and $\overline{\mathrm{IC}}_i$ is satisfied in the region $p_i \le M_{-i} - p_{-i}$ for $i = 1,2$. Next, observe that

$$\begin{aligned} \overline{\mathrm{IR}}_1 \;&\Leftrightarrow\; p_2 \ge \frac{\sigma_2}{\gamma} p_1, \\ \overline{\mathrm{IR}}_2 \;&\Leftrightarrow\; p_2 \le \frac{\gamma}{\sigma_1} p_1. \end{aligned}$$

For the individually rational region to be nonempty, one needs $\sqrt{\lambda_1\lambda_2} \ge \frac{c-b}{b}$, which is satisfied by the assumption $\lambda_i \ge \frac{c-b}{b}$ for $i = 1,2$.

The intersection of $\overline{\mathrm{IC}}_1$ and $\overline{\mathrm{IC}}_1$, using the form those take in the region $\{(p_1,p_2) \mid p_2 \ge M_1 - p_1,\ p_1 \ge M_2 - p_2\}$, is given by

$$p_1 = \frac{M_1 - \sigma_1 M_2}{1 - \sigma_1\sigma_2}, \quad p_2 = \frac{M_2 - \sigma_2 M_1}{1 - \sigma_1\sigma_2}.$$

This intersection occurs above IR_1 if, plugging p_1 above into $\overline{\mathrm{IR}}_1$, we have

$$\frac{M_2 - \sigma_2 M_1}{1 - \sigma_1\sigma_2} \ge \frac{\sigma_2}{\gamma}\frac{M_1 - \sigma_1 M_2}{1 - \sigma_1\sigma_2},$$

or when $\frac{M_1}{M_2} \le \frac{\gamma + \sigma_1\sigma_2}{\sigma_2(\gamma+1)}$; and is below $\overline{\mathrm{IR}}_1$ otherwise. Similarly, the intersection occurs below $\overline{\mathrm{IR}}_2$ if

$$\frac{M_2 - \sigma_2 M_1}{1 - \sigma_1\sigma_2} \le \frac{\gamma}{\sigma_1}\frac{M_1 - \sigma_1 M_2}{1 - \sigma_1\sigma_2},$$

or when $\frac{M_1}{M_2} \ge \frac{(1+\gamma)\sigma_1}{\sigma_1\sigma_2 + \gamma}$; and is above $\overline{\mathrm{IR}}_2$ otherwise.

The slope of $\overline{\mathrm{IC}}_1$ when it binds is $-\sigma_2$ and the slope of $\overline{\mathrm{IC}}_2$ when it binds is $-\frac{1}{\sigma_1}$. The social objective takes the form

$$(b-c)\big(\frac{\sigma_1 - \gamma}{\sigma_1} p_1 + \frac{\sigma_2 - \gamma}{\sigma_2} p_2\big)$$

and has slope $-\frac{\sigma_2}{\sigma_1}\frac{\sigma_1 - \gamma}{\sigma_2 - \gamma}$. The solution is then obtained by comparing slopes in each case. □

References

Adams, Scott, *The Dilbert Principle: A Cubicle's-Eye View of Bosses, Meetings, Management Fads & Other Workplace Afflictions*, New York, NY: HarperCollins Publishers, Inc, 1996.

Al-Najjar, Nabil I., Luca Anderlini, and Leonardo Felli, "Undescribable events," *Review of Economic Studies*, 2006, *73* (4), 849–868.

Aliprantis, Charalambos D. and Kim C. Border, *Infinite Dimensional Analysis*, New York, NY: Springer, 2006.

Anderlini, Luca and Leonardo Felli, "Describability and agency problems," *European Economic Review*, January 1998, *42* (1), 35–59.

Baddeley, Alan, "Working memory: Looking back and looking forward," *Nature Reviews: Neuroscience*, 2003, *4*, 829–839.

Baker, George P., Michael C. Jensen, and Kevin J. Murphy, "Compensation and incentives: Practice vs. Theory," *Journal of Finance*, July 1987, *43* (3), 593–616.

Battigalli, Pierpaolo and Giovanni Maggi, "Rigidity, discretion, and the costs of writing contracts," *American Economic Review*, September 2002, *92* (4), 798–817.

Benabou, Roland, "Groupthink and ideology," *Working Paper*, 2008.

——— **and Jean Tirole**, "Self-confidence and personal motivation," *Quarterly Journal of Economics*, 2002, *117*, 871–915.

Bolton, Patrick and Antoince Faure-Grimaud, "Satisficing contracts," *Review of Economic Studies*, Forthcoming.

Border, Kim C. and Joel Sobel, "Samurai accountant: A theory of auditing and plunder," *The Review of Economic Studies*, October 1987, *54* (4), 525–540.

Cole, Harold L. and Narayana R. Kocherlakota, "Finite memory and imperfect monitoring," *Games and Economic Behavior*, 2005, *53*, 59–72.

Compte, Olivier and Andrew Postlewaite, "Repeated relationships with limits on information processing," *Working Paper*, 2008.

Cowan, Nelson, "The magical number 4 in short-term memory: A reconsideration of mental storage capacity," *Behavioral and Brain Sciences*, 2000, *24*, 87–185.

Dismukes, Key and Jessica Nowinski, "Prospective memory, concurrent task management, and pilot error," in Arthur Kramer, Douglas Wiegmann, and Alex Kirlik, eds., *Attention: From Theory to Practice*, New York: Oxford, 2007.

Dow, James, "Search decisions with limited memory," *Review of Economic Studies*, 1991, *58*, 1–14.

Ericsson, K. Anders and Walter Kintsch, "Long-term working memory," *Psychological Review*, 1995, *102*, 211–245.

Fishburn, Peter C., "Moment-Preserving Shifts and Stochastic Dominance," *Mathematics of Operations Research*, 1982, *7*, 629–634.

Grossman, Sanford and Oliver Hart, "An analysis of the principal-agent problem," *Econometrica*, 1983, *51*, 7–45.

Haseman, W. David, Derek L. Nazareth, and Souren Paul, "Implementation of a group decision support system utilizing collective memory," *Information and Management*, 2005, *42*, 591–605.

Hirshleifer, David and Ivo Welch, "An economic approach to the psychology of change: Amnesia, inertia, and impulsiveness," *Working Paper*, 2001.

Holmström, Bengt, "Moral hazard and observability," *RAND Journal of Economics*, Spring 1979, *10* (1), 74–91.

____ , "Moral hazard in teams," *Bell Journal of Economics*, 1982, *13* (2), 324–340.

Johnson, Norman L. and Samuel Kotz, "Some distributions arising as a consequence of errors in inspection," *Naval Research Logistics Quarterly*, 1985, *32*, 35–43.

Lazear, Edward P., "The Peter Principle: A theory of decline," *The Journal of Political Economy*, 2004, *112*, S141–S163.

____ **and Kathryn L. Shaw**, "Personnel economics: The Economist's view of human resources," *Journal of Economic Perspectives*, Fall 2007, *21* (4), 91–114.

Maskin, Eric and Jean Tirole, "Unforeseen contingencies and incomplete contracts," *Review of Economic Studies*, January 1999, *66* (1), 83–114.

Miller, George A., "The magical number seven, plus or minus two: Some limits on our capacity for processing information," *The Psychological Review*, 1956, *63*, 81–97.

Mirrlees, James A., "Notes on welfare economics, information and uncertainty," in M. Balch, D. McFadden, and S. Wu, eds., *Essays in Equilibrium Behavior under Uncertainty*, Amsterdam: North Holland, 1974.

____ , "The theory of moral hazard and unobservable behavior: Part I," *The Review of Economic Studies*, 1999, *66*, 3–21.

Mohammed, Susan and Brad C. Dumville, "Team mental models in a team knowledge framework: Expanding theory and measurement across disciplinary boundaries," *Journal of Organizational Behavior*, 2001, *22*, 89–106.

Mookherhjee, Dilip and Ivan Png, "Optimal auditing, insurance, and redistribution," *The Quarterly Journal of Economics*, May 1989, *104* (2), 399–415.

Mullainathan, Sendhil, "A memory-based model of bounded rationality," *Quarterly Journal of Economics*, 2002, *117*, 735–774.

Oyer, Paul and Scott Schaefer, "Why do some firms give stock options to all employees?: An empirical examination of alternative theories," *Journal of Financial Economics*, 2005, *76*, 99–133.

Palfrey, Thomas R. and Howard Rosenthal, "Participation and the provision of discrete public goods: A strategic analysis," *Journal of Public Economics*, 1984, *24* (2), 171–193.

Peter, Laurence J. and Raymond Hull, *The Peter Principle: Why Things Always Go Wrong*, New York, New York: William Morrow and Company, 1969.

Piccione, Michele and Ariel Rubinstein, "Finite automata play a repeated extensive game," *Journal of Economic Theory*, 1993, pp. 160–168.

___ **and** ___ , "On the interpretation of decision problems with imperfect recall," *Games and Economic Behavior*, 1997, pp. 3–24.

Rahman, David, "But who will monitor the monitor?," September 2009. Working paper.

Romero, Julian, "Bounded rationality in repeated games," January 2010. Working paper.

Rubinstein, Ariel, *Modeling Bounded Rationality*, Cambridge, Massachusetts: The MIT Press, 1998.

Segal, Ilya, "Complexity and renegotiation: A foundation for incomplete contracts," *Review of Economic Studies*, January 1999, *66* (1), 57–82.

Smith, Rebekah E., "The cost of remembering to remember in event-based prospective memory: Investigating the capacity demands of delayed intention performance," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2003, *29* (3), 247–361.

Stefanski, Leonard, "Monotone likelihood ratio of a 'faulty-inspection' distribution," *The American Statistician*, 1992, *46*, 110–114.

Tirole, Jean, "Cognition and incomplete contracts," *American Economic Review*, March 2009, *99* (1), 265–294.

Van Zandt, Timothy and Xavier Vives, "Monotone equilibria in Bayesian games of strategic complementarities," *Journal of Economic Theory*, May 2007, *134* (1), 339–360.

Wegner, Daniel M., "Transactive memory: A contemporary analysis of the group mind," in Mullen IB and Goethals GR, eds., *Theories of Group Behavior*, New York: Springer-Verlag, 1987, pp. 185–208.

Weldon, Mary S. and Krystal D. Bellinger, "Collective memory: Collaborative and individual processes in remembering," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1997, *23*, 1160–1175.

Williamson, Stephen D., "Costly monitoring, loan contracts, and equilibrium credit rationing," *The Quarterly Journal of Economics*, 1987, *102* (1), 135–145.

Wilson, Andrea, "Bounded memory and biases in information processing," October 2004. Working Paper.

Xiao, Yan, Jacqueline Moss, Colin F. Mackenzie, F. Jacob Seagull, and Samer Faraj, "Transactive responsibility systems and high reliability teams: A tentative formulation," *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting*, 2002.